7-12-2010

N400 Activations in Adults who Stutter in a Picture-Word Priming Task Requiring Attention to Probe Word Phonology

Angela A. Pizon-Moore
University of South Florida

N400 Activations in Adults who Stutter in a Picture-Word Priming Task Requiring Attention to Probe Word Phonology

by

Angela A. Pizon-Moore

A thesis submitted in partial fulfillment
of the requirements for the degree of:
Master of Science
Department of Communication Sciences and Disorders
College of Behavioral and Community Science
University of South Florida

Major Professor: Nathan D. Maxfield, Ph.D.
Stefan A. Frisch, Ph.D.
Joseph L. Constantine, Ph.D.

Date of Approval:
07/12/2010

Keywords: Stuttering, Psycholinguistic, Event-related potentials, Cognitive Neuroscience, Adults

Dedication

This thesis is dedicated to my husband, who not only encouraged me to go back to school, but supported me while I figured out “what I want to be when I grow up”. To my parents, who provided me a foundation education and prayed I would finish my degree. To the lifelong friends I have made while in Graduate School, because of your shoulders to cry on and rallying spirits, I found the strength to complete this process.

To USF’s wonderful faculty of the CSD department, particularly Dr. Maxfield, without whom I would not have been able to complete this thesis. To Dr. Ford, who taught me so much about what it means to be a great mentor and teacher. You both have changed me in ways that will only be discovered as time progresses. I promise to pay it forward!

Acknowledgements

This thesis project was supported in part by funds from the University of South Florida New Faculty Investigator Award, awarded to Nathan Maxfield, PhD.

Table of Contents

List of Tables

List of Figures

N400 Activations in Adults who Stutter in a Picture-Word Priming Task Requiring Attention to Probe Word Phonology

Angela A. Pizon-Moore

Abstract

Objective: A neuroscientific picture-word task was used to investigate semantic and phonological activation spreading in adults who stutter (AWS).

Method: Fourteen AWS and 14 adults who do not stutter (AWNS) participated. On each trial, a picture was named at a delay. Sometimes, an attended probe word was heard before naming. Some probes were Semantically-Related to the labels. Those same probes also appeared following pictures with Unrelated labels. N400 ERPs recorded to these two probe types were compared (Semantically-Related versus Unrelated). Other probes were Phonologically-Related to the labels. Those same probes also appeared following pictures with unrelated labels (P-Unrelated). N400 ERPs recorded to these two probe types were compared (Phonologically-Related versus P-Unrelated).

Results: AWNS exhibited typical N400 priming effects. AWS exhibited non-robust Semantic N400 priming, and a reverse Phonological N400 priming effect.

Conclusions: Results suggest that AWS use attentional control strategies to influence the activation of words in the mental lexicon.

Introduction

Approximately 2.4% of all people stutter chronically at some point during their lives (Craig, Hancock, Tran, Craig, & Peters, 2002). Stuttering is typically diagnosed during childhood, but only in ~20% of cases does it persist into adulthood (Craig & Tran, 2005), affecting ~1% of adults or ~4 million adults in the U.S. (Craig et al., 2002; Craig & Tran, 2005). More men are affected than women at a ratio of 2.3:1 (Craig et al., 2002; Craig & Tran, 2005).

The impact of persistent stuttering on quality of life has been considered by Yaruss (2007) in the context of the World Health Organization's model of International Classification of Functioning, Disability and Health (ICF). As shown in Figure 1, the ICF model considers the impact of a disorder on each of three areas: Body function and structure, Activities, and Participation. "Activities" include, among other things, one's ability to communicate in his or her current environment (O'Halloran & Larkins, 2009). At this level of functioning, stuttering can impact one's ability to acquire a place to live, utilize goods and services, and achieve economic self-sufficiency (Yaruss, 2007). "Participation", in the ICF system, refers to the social roles an individual assumes, and what those require in terms of social interaction (O'Halloran & Larkins, 2009). Here, stuttering can affect one's ability to start, maintain or end conversations, establish and maintain interpersonal relationships, interact according to social rules, and participate in social events (Yaruss, 2007).

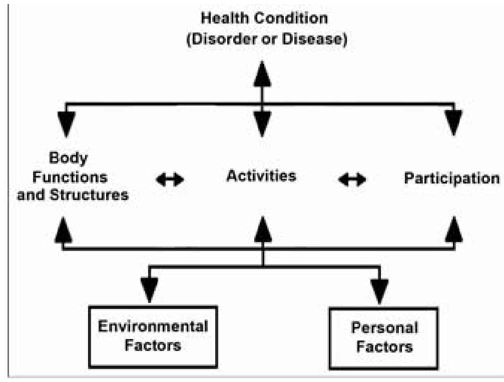

Figure 1. World Health Organization (WHO) International Classification of Functioning (2001). *WHO. International Classification of Functioning, Disability, and Health*, p. 18.

Stuttering can also affect one's body functions and structures, the third function addressed in the ICF model. Voice production, articulation, and fluency may all be affected with stuttering. Stuttering can affect emotional functioning, too. For example, some adults who stutter (AWS) - in particular, those who experience bullying and teasing as children - demonstrate a disproportionately high rate of generalized anxiety disorder (Craig, Blumgart, & Tran, 2009; Hugh-Jones & Smith, 1999; Davis, Howell, & Cooke, 2002).

Having outlined some of the effects that stuttering can have on a person; we now operationally define the disorder. A common definition of stuttering is as follows (Wingate, 1964, p. 488):

> *Disruption in the fluency of verbal expression, which is characterized by involuntary, audible or silent, repetitions or prolongations in the utterance of short speech segments, namely: sounds, syllables, and words of one syllable. These disruptions usually occur frequently or are marked in character and are not readily controllable. Sometimes the disruptions are accompanied by accessory activities involving the speech apparatus, related or unrelated body structures, or stereotyped speech utterances. These activities give the appearance of being a*

speech-related struggle. Also, there are not infrequently indications or reports of the presence of an emotional state, ranging from a general condition of "excitement" or "tension" to more specific emotions of a negative nature such as fear, embarrassment, irritation, or the like.

This definition focuses largely on motor manifestations of stuttering. In more than four decades of research, it was established that speech motor mechanisms operate differently in AWS versus adults who do not stutter (AWNS; see overview by Peters, Hulstijn, & Van Lieshout, 2000). Perhaps as a result of this focus on motor performance in AWS, many treatments for stuttering emphasize speech motor performance; an approach that has been shown to be efficacious (Conture, 1996; Bothe, Davidow, Bramlett, & Ingham, 2006). Still, AWS relapse at rates estimated between 30% and 50% (Boberg & Kully, 1994; Howie, Tanner, & Andrews, 1981; Perkins, 1981), and as high as 75% (Craig & Hancock, 1995). Relapse is probably not the result of a single factor; rather, a combination of factors probably drives relapse including stuttering severity, attitude, personality type, locus of control, and motivation (Craig, 1998).

Another factor that potentially contributes to stuttering relapse is that deficits in the core mechanisms of speech production are not comprehensively addressed as part of treatment. Speech production is driven by motor, psycholinguistic, and non-linguistic cognitive mechanisms. In order to produce speech fluently, all of the mechanisms involved must operate efficiently. The aim of this study was to further our understanding of psycholinguistic functioning in AWS; specifically, of lexical activation spreading, a process by which words become activated in the mental lexicon on the path to speaking.

In the section below, we outline a model of speech production mechanisms, including psycholinguistic mechanisms, with an emphasis on lexical activation spreading. We then review models of stuttering etiology, including the Covert Repair Hypothesis, which implicates inefficient activation spreading as contributing to moments of stuttering. Next, we review the current evidence on activation spreading operations in AWS, highlighting some of the shortcomings in our current knowledge. Finally, we outline a neuroscientific approach to studying lexical activation spreading in AWS; an approach that addresses some of the methodological shortcomings of past research.

Mechanisms of Speech Production: An Overview

Human beings speak as a means of mediating their environments, exchanging information, and gaining understanding (Wertsch, 1993). The intention to speak triggers a set of processing events and mechanisms. As outlined in Levelt's (1989) "blueprint" model of speech production (see Figure 2), a message is conceptualized, formulated, verbalized, and checked for integrity via a complex set of psycholinguistic, motor, and cognitive mechanisms; all of which must operate efficiently in order for speech to be produced fluently.

Message generation begins with conceptualization. Various components of the speaker's intended message are retrieved from long term memory and deposited into working memory, including the topic and focus of the message. Experiential and situational concepts, also retrieved from long term memory, are pieced-together, creating crude message fragments; each conveying parts of the speaker's intended message.

Next, the pre-verbal message fragments are specified linguistically, i.e., translated into meaning-carrying utterances; a process termed formulation in Levelt's model (see Figure 2). Lemmas (i.e., meaning-carrying, or content, words) matching concepts in the preverbal message are activated in, and retrieved from, the mental lexicon. As lemmas become activated, their semantic and grammatical-category properties become available. In connected speech, a grammatical encoding mechanism generates syntactic structure via linguistic constraints, assigning each lemma to a specific function and position within the syntactic structure. Each lexically- and syntactically-specified utterance is subsequently stored in the syntactic buffer for phonological encoding.

During phonological encoding, the articulatory plan for each utterance is specified. The segment (phoneme) string for each word comprising the utterance is retrieved. Separately, the metrical properties of each word (i.e., its syllable structure and stress pattern) are accessed. The segmental and metrical features for each of the words comprising an utterance are combined to create phonological words, in which adjustments are made for co-articulation (Levelt & Meyer, 2000).

According to the sensorimotor control model by Van der Merwe (1997), the phonological plan is translated into movement patterns in two steps. First, a speech motor plan for each utterance is created that specifies a set of temporal and spatial goals for articulatory movement. Articulatory context is monitored, and the motor plan is adapted for co-articulation, improving efficiency and simplifying the plan directives. A speech motor program is then generated, which specifies the kinematics (e.g., trajectory of movement) and force profile (i.e., amount of velocity, range, and force needed for each movement). The transmitted motor program not only drives the speech muscles; it also provides a template for tactile-proprioceptive and auditory self-monitoring. Guenther (2006) proposed that the compiled motor program is sent not only to the primary motor strip, but also to neuroanatomical regions subserving self-monitoring. According to Guenther (2006), proprioceptive information is received by the primary sensory strip and then compared with the original program; a process Guenther (2006) suggests is mediated by supramarginal gyrus. In addition, auditory information received by primary auditory cortex is compared with the original program stored; a process Guenther (2006) suggests is mediated by superior temporal gyrus. Errors detected externally, through these two channels, trigger correction.

According to Levelt's (1989) model, linguistic monitoring occurs, too, both internally and externally. Internally, the emerging linguistic plan is 'checked' for cohesiveness and accuracy against the lemmas and preverbal message, before it is sent forward for speech motor planning and programming. Externally, the form, content and well-formedness of the utterance is checked.

If linguistic errors are detected, it is believed that an error-repair process is triggered while the utterance is halted (see Levelt, 1983). It is through this internal monitoring and repair cycle that errors in the psycholinguistic formulation of utterances - specifically in the activation and retrieval of words - are believed to result in both normal and stuttering-like disfluencies.

A focus on lexical activation spreading. The average adult lexicon contains upwards of 50,000 to 100,000 words (Miller, 1991). During vocabulary acquisition, each word is "filed" in the mental lexicon and organized. Picture naming provides a simplified model for understanding how words become activated in, and retrieved from, the mental lexicon during speech production (Glaser, 1992). In the fraction of a second that separates the onset of presentation of a picture from the articulation of its label, a network of words whose meanings relate to the pictured object are activated in the mental lexicon of the speaker. We refer to this as semantic activation spreading. The lexicon is also organized phonologically, and each initially-activated word activates still other words sharing the same phonemes. We refer to this as phonological activation spreading. The set of semantically-related picture labels, and their phonologically-associated words, competes for activation (see Dell, 1986). Competition for selection occurs and, ideally, a target word emerges.

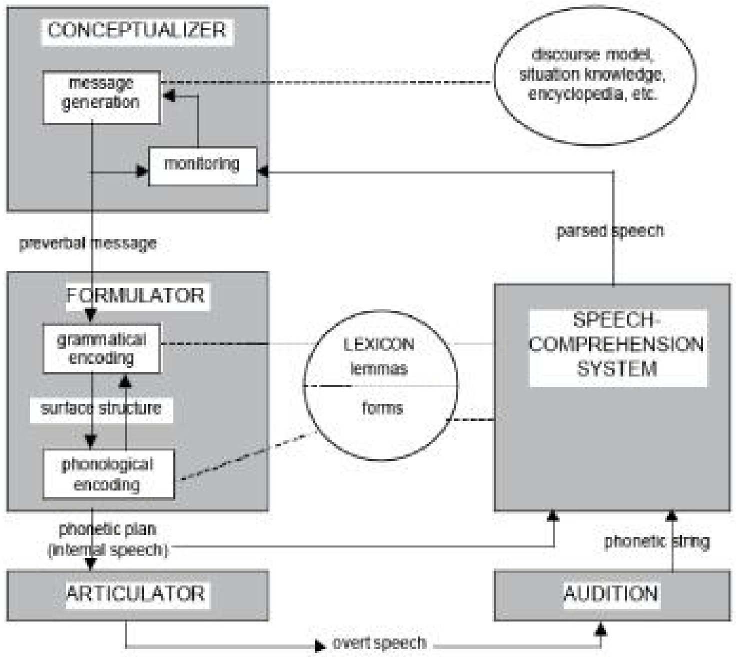

Figure 2. A blueprint of speech production. Levelt, (1989). *Speaking: From intention to articulation,* Published by MIT Press, p.9. Reprinted with permission.

A Psycholinguistic Model of Stuttering

According to at least one hypothesis, psycholinguistic breakdown occurs in AWS because a clear "winner" in the competition for activation among words does not always emerge (Postma & Kolk, 1993), with the following undesirable consequences. At some point during psycholinguistic planning, the system triggers, i.e., forces, the retrieval of phonemes for the word with greatest activation strength (Kolk & Postma, 1997). If competition among words is unresolved, then the phonemes of a strongly activated yet unintended word may be retrieved in error. An internal monitor, which checks the surface structure of psycholinguistic plans for accuracy (Levelt, 1983), may then detect the presence of ill-retrieved phonemes and signal the speaker to halt speaking and initiate a repair; potentially disrupting fluency (Dell, 1986; Kolk, & Postma 1993, 1997).

This theory, known as the Covert Repair Hypothesis (Kolk & Postma, 1993), is based on the work of Hockett (1967, 1973) and Levelt (1983), who studied monitoring and self-repairs of speech errors in typically fluent speakers. The term "covert editing" was coined by Hockett (1967, 1973) to explain how prearticulatory editing might occur. According to his view, during the formulation step - when the phonological and motor movement plan are retrieved and integrated – this information is sent to an internal monitor that detects errors and sends the package for reprocessing. Reprocessing occurs prior to initiation of speech if there is time and thus, the error is undetected by the listener. If overt speech has already begun when an error is detected, the message is sent for reprocessing. However, it is believed that AWS still attempt to produce the incorrect plan again and again as a stall tactic while the revision to the phonetic plan occurs (Postma & Kolk, 1993). Time pressure during speech production is believed to further exacerbate this cycle of repair-driven disfluency (Perkins, Kent, & Curlee, 1991).

Current evidence on semantic activation spreading in AWS*.* Efficient activation spreading occurs as the result of two factors: One is the formation of appropriate connections between related words. Second, limits must be placed on the number of related words able to compete at any one time (Dell & O'Seaghdha, 1991, 1992). Some evidence suggests that semantically-related words are poorly networked in AWS. For example, AWS were shown to score lower than AWNS on the Peabody Picture Vocabulary Test (Prins, Main, & Wampler, 1997), a measure of polysemy (Miller & Lee, 1993). Other evidence suggests that while competition among words closely-related in meaning is resolved typically in AWS (Hennessey, Nang, & Beilby, 2008), words whose meanings are not optimally-suited for expressing target concepts can undesirably accrue activation. For example, AWS were shown to produce less common word associations than AWNS (Wingate, 1988), and more confrontation naming errors (Newman & Ratner, 2007). AWS scored lower than AWNS on the Verbal Scale of the Wechsler Adult Intelligence Scale, defining words more verbosely than AWNS while using fewer synonyms (Wingate, 1988). AWS were also shown to have more difficulty than AWNS disambiguating word meanings in confusing sentence contexts (Watson et al., 1994). This evidence points to a potential weakness in how words become activated according to their semantic features, and compete, in AWS.

Current evidence on phonological activation spreading in AWS. There is limited evidence, too, about the operation of phonological activation spreading in AWS. Much of the existing evidence is rooted in reaction time measures. In one study, AWS exhibited similar speech reaction times with phonological priming versus AWNS (Burger & Wijnen, 1999; Hennesey, et al., 2008), suggesting that phonological activation in AWS is similar to AWNS with typical competition resolution. However, another group of studies paints a different picture. In one study, AWS were slower than AWNS in recognizing pairs of orthographically similar but phonologically different words (Weber-Fox, Spencer, Spruill, & Smith, 2004). During tongue twisters (Postma, Kolk, & Povel, 1990; Eldridge & Felsenfed, 1998) and moments of stuttering (Viswanath, Poindexter, & Rosenfeld, 1999), AWS exhibited increased sound production errors. In another study by Dell (1990), high rates of stuttering were elicited by low frequency words, suggesting interference from high frequency phonologically-related neighbors. This body of evidence tentatively suggests deficits in phonological activation spreading, too, in AWS.

A Neuroscientific Approach to the Study of Activation Spreading in AWS

As outlined above, most of the psycholinguistic research conducted to date with AWS was driven by tests of naming, lexical decision, word association, vocabulary knowledge, and changes in stuttering frequency as a function of linguistic context. This approach, while providing useful insight into psycholinguistic planning in AWS, is limited in its ability to describe this process. AWS tend not to produce frequent, overt linguistic errors that are used to diagnose psycholinguistic deficits in other groups, e.g., people with aphasia (Kay, Coltheart, & Lesser, 1992). Furthermore, behavioral measures are not synched closely enough to specific psycholinguistic mechanisms to detect specific deficits in them (Hagoort & Kutas, 1995). In contrast, neuroscience measures, which are sensitive to cognitive processing, have proven fruitful in refining descriptions of cognitive-linguistic deficits, e.g., in aphasia (Caramazza & Coltheart, 2006). The aim of this study was to extend behavioral evidence on activation spreading in AWS, reviewed above, by describing this process with greater precision, using event-related potentials (ERPs).

ERPs are voltages recorded at the scalp that reveal at least some of the electrical activity generated by the brain. ERPs activate reflexively as people process stimuli, make decisions, and regulate their behavior (Otten & Rugg, 2005). Crucially, the averaged ERP signal is comprised of many different components, each indexing the activation of a distinct cognitive process.

Of primary interest here was the N400, an ERP component defined by negative-going activity that peaks in amplitude at ~ 400 to 550 ms after the onset of a stimulus, usually a word (Fischler, 1990). The amplitude of the N400 component has been shown to be inversely related to the strength of activation that emerges from a priming context (Van Petten & Kutas, 1991; Rosler & Hahne, 1992). That is, a word whose activation has been primed by a preceding stimulus will elicit N400 activity that is smaller in amplitude than when the same word has not been primed by a preceding stimulus.

Weber-Fox and colleagues have used these properties of the N400 to assess lexical activation in AWS. In one study (Weber-Fox, 2001), participants read sentences silently, some of which contained word violations, e.g., She looked at her watch to check the *rain*. N400, while expectedly large for AWNS in response to the semantically-incongruous words, was reduced in amplitude for AWS. Weber-Fox and Hampton (2008) reported similar results from an auditory sentence task. While such findings underscore the processing-sensitivity of ERPs, the extent to which atypical sentence processing mirrors psycholinguistic planning is controversial (Frazier, 1982; Garrett, 1988; Kempen, 2000). For example, there is evidence that different input and output lexicons exist

(Monsell, 1987; Cutting, 1998), and it may be difficult to know what ERPs elicited by words on input reveal about how words are processed in preparation for speaking.

N400 priming can also be used to assess activation spreading on the path to picture naming, using a method called picture-word priming (see Jescheniak, Schriefers, Garrett, & Friederici, 2002). Picture-word priming involves presenting a picture on each trial, followed by an auditory probe word. Participants are instructed to label the picture, but not to name it aloud until several hundred milliseconds (ms.) after the probe word has been presented (i.e., when prompted by a cue to name the picture). Each probe word elicits ERP activity, including activation of the N400 component.

Jescheniak et al. (2002) hypothesized that when preceding picture label and subsequent probe word are unrelated, N400 activated in response to the probe word should be relatively large in amplitude. If, on the other hand, the preceding picture label and subsequent probe word are related in some way, N400 amplitude should be attenuated. The authors confirmed these effects, reporting that both semantic picture-word priming (e.g., picture of *grass*, followed by probe word *mower*) and phonological picture-word priming (e.g., picture of *grass*, followed by probe word *grab*) attenuated N400 amplitude in AWNS when they were required to attend to and remember the probe words.

A version of this task was recently used to investigate semantic activation spreading in AWS (Maxfield, Huffman, Frisch & Hinckley, in press). In that study, phonological rehearsal of the probe words was deemphasized. For the AWNS control group, N400 amplitude was reduced when probe words were preceded by a picture with a semantically-related label relative to when probe words were preceded by a picture with a semantically-unrelated label. In contrast, for the AWS group, N400 amplitude was enhanced when probe words were preceded by a picture with a semantically-related label relative to no priming. Maxfield et al. (in press) gave three different interpretations of this effect. One interpretation was that the mental lexicons of AWS are poorly organized, resulting in weak activation of a target picture label and the triggering of an inhibitory mechanism that allowed the target label to accrue greater activation strength than its semantically-related neighbors. The second interpretation of the N400 reversal seen in AWS was that "too many" semantically-related words become activated on the path to naming due to the tendency for AWS to use word substitutions and circumlocutions; again requiring an inhibitory mechanism to allow the target label to emerge. A third interpretation was that, because at least some AWS have reduced phonological working memory capacity (Bajaj, 2007), holding a picture label in mind while a semantically-related probe word is presented might be distracting enough to require semantic inhibition. In the current study, we required attention to the auditory probe words to investigate what effect phonological rehearsal might have on processing in AWS given the effects seen in Maxfield et al. (in press).

Summary and Research Questions

Stuttering affects the lives of AWS in many facets, including social and vocational. Past research focused primarily on motor contributions to stuttering. For more than two decades, research in the area of stuttering has shifted its focus to higher-level processes involved in speech production, too. However, little progress has been made; in part due to methodological limitations.

We used a picture-word N400 priming task developed by Jescheniak et al. (2002) to answer two research questions: 1) Does semantic activation spreading operate typically in AWS, as evidenced in the amplitude of the N400 ERP component elicited during a picture-word priming task? 2) Does phonological activation spreading operate typically in AWS, as evidenced in the amplitude of the N400 ERP component elicited during a picture-word priming task? If activation spreading in AWS works typically, then

N400 should attenuate in amplitude with semantic or phonological priming, as seen for AWNS.

Method

Participants

A total of 34 individuals were tested including 15 AWS and 19 AWNS. Of the 15 AWS who participated, data for 14 of those AWS were analyzed (11 males and 3 females, with a mean age of 32.2 years, ranging from 20 to 52 years old). One AWS was excluded due to excessive EEG noise. Of the 19 AWNS who participated, data for 14 AWNS were analyzed (2 males and 12 females, with a mean age of 23.89 years, ranging from 19 to 40 years old). Five AWNS were excluded because they had speech therapy as children (n=2), exposure to another language at a young age (n=2), or excessive EEG noise (n=1). We note that, although the ratio of women to men was nearly reversed between groups, our analysis of the ERP data – reported below – did not evidence differences between groups indicative of gender (i.e., different patterns of lateralization in the ERP components, see Van Dyke, Zuverza, Hill, Miller, Rapport, & Whitman, 2010).

All included participants were monolingual speakers of English with normal or corrected-to-normal vision, no hearing deficit, and normal language function. None of the AWNS were taking medications that can affect cognitive function, and none had a history of neurological injury. One of our AWS reported taking a prescription anti-anxiety medication at time of testing. None of the AWS participants had a history of neurological injury. A speech sample was collected from each of the 14 AWS, and analyzed by a Speech-Language Pathologist with a Certificate of Clinical Competency from the American Speech-Language-Hearing Association, in order to confirm their diagnosis of stuttering. All participants gave informed consent to participate in the study, completed a medical history questionnaire, and were compensated 10 dollars per hour.

Stimuli

The stimuli for this study consisted of 54 simple line drawings of culturally common objects, selected from the International Picture Naming Project (IPNP) Mini Database Query, a database of pictures whose visual quality and labels have been normed (Szekely et al., 2004). All objects were depicted as black-line drawings having similar style and quality. Each image measured 2 inches in height by 2 inches in width. The most frequently-used label for each drawing, as determined by the IPNP norms, was a noun. The picture labels were either 1 or 2 syllables in length with an average number of phonemes of 3.89. Figure 3 shows examples of four picture stimuli used in the study.

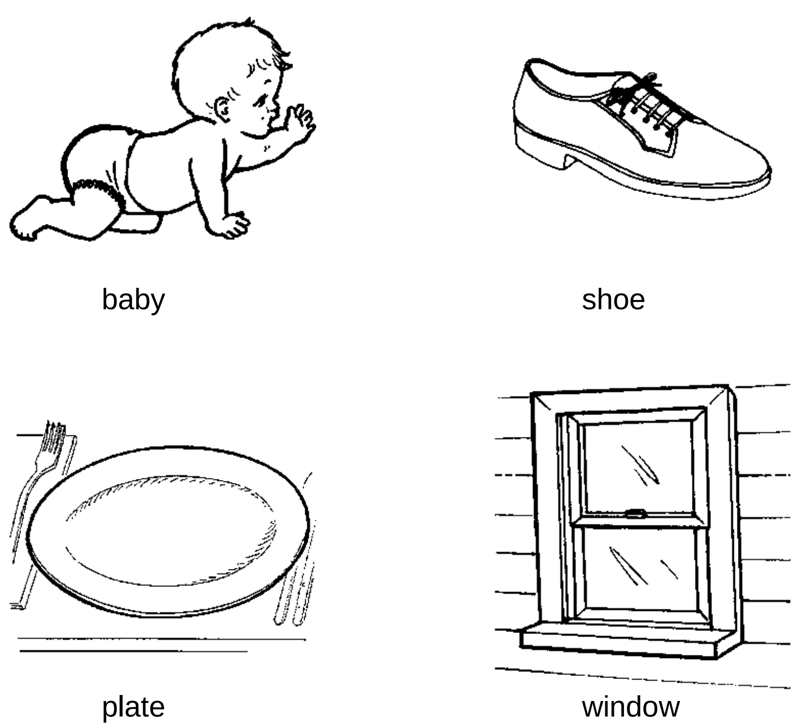

Figure 3. Examples of four black-line drawings used in the study, each shown with its label.

Two auditory probe words were assigned to each picture. One was the strongest semantic free associate of the picture label, as determined using the University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 1998). These Semantically-Related probe words were phonologically unrelated to their corresponding picture labels. The second probe word chosen for each picture contained the same initial consonant-vowel (CV) onset combination as the picture's label, i.e., they were strongly Phonologically-Related to the corresponding picture label. Words were not included in the Phonologically-Related probe set if a semantic relationship existed between the word and corresponding picture label. Words were not included into either set if second syllable stress was present.

Each of the two probe words selected for each picture was reassigned to a different picture with no relationship. For instance, the picture "shoe" (see Figure 1) was assigned the phonologically-related word "shoots" (creating a Phonologically-Related condition) and the semantically-related word "foot" (creating a Semantically-Related condition). The same words, "shoots" and "foot", were then each reassigned to another picture whose label was completely unrelated in form or meaning (when Phonologically-Related words were reassigned to a completely unrelated picture, we labeled them P-Unrelated; when Semantically-Related words were reassigned to a completely unrelated picture, we labeled them Unrelated). Thus, each probe word appeared twice: once in a Related condition, and once in an Unrelated condition. Appendix A lists each picture label with its assigned probe words. All of the words in each set were checked on a measure of word familiarity in the Hoosier Mental Lexicon (Nusbaum, Pisoni, & Davis, 1984). The Hoosier Mental Lexicon is a collection of the pronunciations and familiarity ratings of approximately 20,000 American English words based on Webster's Seventh Collegiate Dictionary (1967). The average familiarity rating for the picture labels was 6.59 (SD=0.1). The average familiarity rating for the semantically related words was 6.98

(SD=.05). The average familiarity rating for the phonologically related words was 6.61 (SD=0.33). High average familiarity ratings suggest that that the picture labels utilized in the experimental and filler trials represent words that are highly familiar to speakers of American English.

In addition to creating the probe words, a list of words was created that were used in a word verification task on Experimental trials of the task. As described in the Procedure below, at the end of each Experimental trial, participants saw a printed word that was either identical to the probe word they had heard or one that rhymed with it. In the latter case, for instance, if the auditory probe word was “stone” the word “phone” might have been presented for the word verification. For the 54 Semantically-Related probe words selected as described above, a total of 27 rhyming words were selected for the word verification while the other 27 were identical to their corresponding probe words. The 27 rhyming words had an average familiarity rating of 6.95 (SD=0.86). Likewise, for the 54 Phonologically-Related probe words selected as described above, a total of 27 rhyming words were selected for the word verification while the other 27 were identical to their corresponding probe words. Those 27 rhyming words had an average familiarity rating of 6.9 (SD=0.148). When the probe words were reassigned to the Unrelated or P-Unrelated conditions, respectively, a different word check list was composed. For most of the cases (n=68), a probe word that had been followed by itself was replaced by a rhyming verification word and vice-versa. For a smaller subset of cases (n=38), a probe word was followed by itself in the word verification both times it appeared. For an even smaller subset of cases (n=2), a probe word was followed by a rhyming verification word both times it appeared. These latter word check conditions were included to prevent participants from predicting when a probe word might be followed by itself versus a rhyming word.

Preparation of Auditory Probe Words

The probe words were transformed into a set of auditory stimuli as follows. A female, native speaker of English read aloud each word, several times consecutively. All readings were recorded to digital audiotape, digitized at a sampling rate of 44.1 kHz, and then processed using Sony Sound Forge 8.0 editing software. The best spoken exemplar of each word was selected and its waveform spliced from the original recording, then saved as a separate sound file (.wav format). The sound files were edited to eliminate periods of silence prior to the word to ensure the auditory stimulus was immediately initiated once the sound file was played. The loudness of each word was normalized to an RMS amplitude of 15 dB, and a noise gate was used to reduce high-frequency noise (hiss).

Procedures

Prior to testing, participants were familiarized with the picture stimuli, making sure they knew the target label for each picture. Instructions indicated that two types of trials would be presented during testing. Participants were told that, on one trial type (Filler trials) they would see a picture and were to name the picture aloud after a naming cue (!!!) appeared on the screen. Participants were told they could proceed to the next trial by pushing any button on a response box. The AWS participants were further instructed that if they stuttered while naming, to finish saying the word completely and then press the button to continue.

Additionally, participants were told that, on the second trial-type (Experimental trials) a picture would appear on the screen, followed by a word presented through the earphones, and that they were to verbally name the picture after a naming cue (!!!) was displayed. They were also told that a question would appear after the naming cue, asking: “Is the word you heard <u>[printed verification word shown here]</u>?” They were instructed to use the response box and press number one for yes and number five for no. Here, too, the AWS participants were told to continue stuttering and finish the word before verifying the probe word on the response box. Before testing began, participants were asked to verbally restate the task instructions using their own words to ensure understanding of the task requirements. Participants were then asked to minimize extraneous movements while participating in the experiment.

Each participant was tested in a single session with a short break roughly halfway through the task. Each participant received a total of 324 trials (54 Phonologically-Related trials, 54 P-Unrelated trials, 54 Semantically-Related trials, 54 Unrelated trials, and 108 Filler trials). As shown in Figure 4, Experimental trials consisted of a crosshair (+) shown for 500 ms (ms), followed by a black-line drawing for 400 ms, followed by another crosshair for 150 ms, followed by an auditory probe word to which ERPs were measured, followed by a naming cue ("!!!") shown for 1000 ms, followed by a word verification that remained on-screen until the participant pressed a button, prompting a new trial. There were 1,350 ms from the onset of the line drawing to the onset of the verbal naming cue.

As shown in Figure 5, Filler trials consisted of a crosshair shown for 500 ms, followed by a black-line drawing for 400 ms, followed by another crosshair shown for 300 ms, and then a naming cue ("!!!") that remained on-screen until the participant named the picture aloud and then pressed a button on the response box to continue. 1450 ms separated the onset of the picture from the onset of the articulation cue, which stayed on the monitor until the participant pressed the button to begin the next trial. There were 700 ms from the onset of the line drawing to the onset of the verbal naming cue.

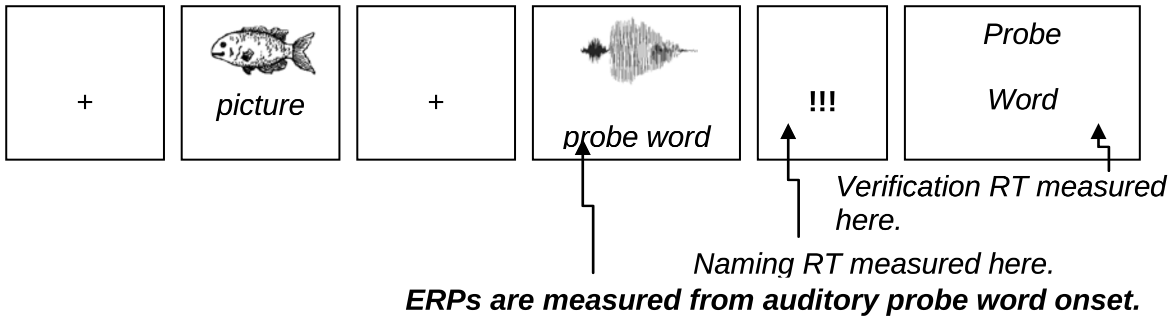

Figure 4. Experimental trial structure.

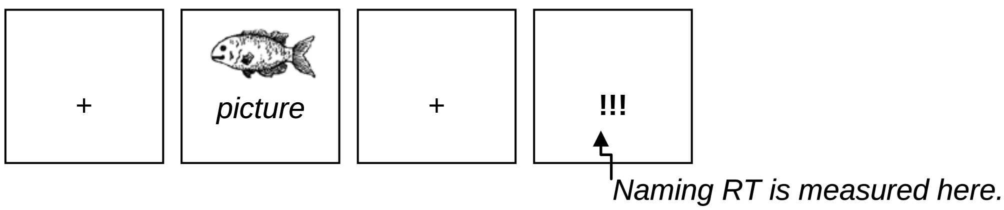

Figure 5. Filler trial structure.

The 324 items were presented in a single, large block of trials with a break at the halfway mark. Trials for each of the five different conditions were presented in random order. Each of the 54 pictures appeared a total of six different times during testing: Twice in Filler trials, and once in each of the four probe word conditions. This procedure is the same as the experimental design used by Jescheniak et al. (2002).

Apparatus and Recording

Each participant was seated in a dimly lit, sound-attenuating booth, facing a 19-inch LCD computer monitor. The spoken probe words were presented auditorily via high-quality, insert earphones (Etymotic Research, Model E-2). Participants signaled the experimental software (Eprime) to progress from one trial to the next by using a push-button response box (Psychological Software Tools). In addition to behavioral data, continuous EEG was recorded from each participant as follows. During testing, each participant wore a nylon QuikCap (Neuroscan). The cap was fitted with a set of 62 active recording electrodes, positioned in a geodesic pattern covering the forehead, top, sides, and back of the head, as well as one reference (midline Cz reference) and one ground electrode. Four additional electrodes recorded electro-ocular activity. A recording electrode was also affixed to each mastoid process. The electrodes were constructed of silver / silver chloride (Ag / AgCl). Conductive electrolyte QuikGel (Neuroscan) was used as the medium between each electrode and the scalp. Placement of the cap took between 10 and 30 minutes.

Continuous EEG was recorded from each participant during testing at a sampling rate of 500 Hz (1 recording every 2 ms from each electrode). SCAN software, Version 4.3 (Neuroscan), was used to control EEG recording. Electrode impedance was kept below 30 kOhm (Ferree, Luu, Russell, & Tucker, 2001). The continuous EEG data were low-pass filtered online, at a corner frequency of 100 Hz. E-Prime experimental control software (Psychological Software Tools, version 1.1), was run on a PC computer and used to present the stimulus pictures.

EEG-to-Average ERP Data Reduction

The continuous EEG record of each participant was segmented into individual epochs. Each epoch was comprised of EEG data recorded from each electrode during presentation of the auditory probe word on each Experimental trial, beginning 300 ms before the onset of the word and terminating 1000 ms following word onset. Epochs of the same duration were also created for each Filler trial, beginning 300 ms before picture onset and terminating 1000 ms later. Although the latter epoch types were not ultimately analyzed, including them beneficially increased the amount of data used as input to an ocular artifact correction procedure, described below. Experimental and Filler trials eliciting incorrect picture names were excluded. We began with an extended epoch length (-300 to 1000 ms) to ensure that artifacts on the leading and trailing edges of the epoch interval we analyzed (-100 to 800 ms) were rejected or corrected. The ERP data were truncated to that shorter interval (-100 to 800 ms) after averaging.

Ocular artifact correction*.* Inspection of the EEG data revealed that most participants' recordings were contaminated by eye blink artifact. In order to salvage as many trials as possible (Picton et al., 2000), we used an ocular artifact correction procedure described in Glass et al. (2004). The segmented, baseline-corrected EEG data for each participant were submitted to an Independent Component Analysis (ICA) (Bell & Sejnowski, 1995). After the ICA decomposition of each EEG record into 64 components, the inverse weights (scalp map) of each component were correlated with a blink template generated by averaging at each channel the peak activity of two blink exemplars sampled from each participant. Any component whose inverse weights matched the blink template (r = .9 or better) was identified as a blink component. The activity related to each blink component was removed from each trial if it reduced the overall EEG variance for that trial. For 12 of the 14 AWS, and for 10 of the 14 AWNS, a blink component was identified. A blink component was not reliably modeled using ICA for two of the 14 AWS, and for 4 of the 14 AWNS. For those six individuals, traditional artifact rejection was used (see EEG trial rejection, below). Of the 22 participants for whom a blink component was reliably modeled, an average of 267 trials (SD=32) were corrected for blink activity.

EEG trial rejection*.* For the six individuals whose data were not corrected for blink activity using the ICA approach outlined above, a traditional artifact rejection procedure was used whereby any trial with activity greater than or equal to 100 microvolts at the VEOG and HEOG leads were rejected. None of those six individuals lost more than one-third of their trials due to ocular artifact.

After ICA blink correction (for n=22 participants) and ocular artifact rejection (for n=6 participants), the data for all 28 participants were further checked for trials with noisy active recording electrodes. Channels whose fast-average amplitude exceeded 200 microvolts (large drift) were marked bad; as were channels whose differential amplitude exceeded 100 microvolts (high-frequency noise). Any EEG trial with more than three bad channels (5% of the total number of channels) was rejected from further analysis. No participant lost more than 20% of their trials for any condition due to bad channel artifact; most lost well under 10% of their trials.

Final EEG processing. For any accepted trial with channels marked bad (<=3), the EEG activity at those channels was replaced using a three-dimensional spline interpolation procedure implemented in Matlab (Nunez & Srinivasan, 2006, see Appendices J1-J3). Accepted EEG trials were then averaged together, separately for each Experimental condition. As a result, each participant had four sets of ERP averages: Semantically-Related, Unrelated, Phonologically-Related, and P-Unrelated. For each participant, no fewer than 34 artifact-free trials went into the set of ERP averages for each Experimental condition. The averaged ERP data were low-pass filtered at a corner frequency of 40 Hz, truncated to the critical time window (-100 to 800 ms), re-referenced to the left mastoid electrode, and baseline-corrected (-100 to 0 ms).

Analysis

Analysis of behavioral data. Naming responses were scored for accuracy. Naming on each trial was judged correct if the participant named the picture using the one-word label indicated during the familiarization phase. Naming was judged as incorrect if no response was given, if a participant began responding before the prompt (!!!) was displayed (coded as early response), if a participant responded after the prompt (!!!) was no longer being viewed (coded as delayed response), a whole-word error was generated (e.g., “stone” for “rock”), or a phonological error was generated (e.g., “thesk” for “desk”). Trials scored as incorrect were excluded from further analysis. The percentage of correct naming trials was computed for each participant in each condition. In order to test for statistically significant effects, the percent correct scores were submitted to a repeated-measures ANOVA with Condition entered as a within-subjects factor having five levels (Filler, Semantically-Related, Unrelated, Phonologically-Related, P-Unrelated) and Group entered as a between-subjects factor having two levels (AWS, AWNS).

Additionally, each participant's word check responses were scored for accuracy automatically, by E-Prime software, as participants gave their button-press responses at the end of each probe word trial. Trials scored as incorrect were also excluded from further analysis. The percentage of correct word check responses was computed for each participant in each probe word condition. In order to test for statistically significant effects, the percent correct scores were submitted to a repeated-measures ANOVA with Condition entered as a within-subjects factor having four levels (Semantically-Related, Unrelated, Phonologically-Related, P-Unrelated) and Group entered as a between-subjects factor having two levels (AWS, AWNS).

For any test violating the assumption of sphericity we report p-values based on adjusted degrees of freedom (Greenhouse & Geiser, 1959) along with original F-values. When a statistically significant main effect or interaction was detected, Bonferroni-protected pair-wise comparisons were made.

Analysis of ERP Data. The aim of this analysis was to demonstrate N400 ERP amplitude differences between conditions (i.e., Semantically-Related versus Unrelated N400 amplitude differences, as well as Phonologically-Related versus P-Unrelated N400 amplitude differences), and to compare the direction and magnitude of these differences between groups. A standard approach to ERP amplitude measures involves visually inspecting a set of grand averages, selecting time windows of interest, and measuring the amplitude of the ERP activity during those time windows, for each participant, in each condition, at specified electrodes. This approach is limited in at least two ways. First, investigator judgment drives the choice of time windows. Second, average amplitude measures do not effectively deal with component overlap (Dien & Frishkoff, 2004).

In order to address these limitations, we used Principal Component Analysis (PCA) - a non-parametric technique that begins with a data set comprised of many dimensions and reduces it to a subset of smaller dimensions by forming linear combinations of the original set of variables. Each linear combination of variables is called a factor. Each factor is defined by a subset of the original variables that are redundant enough (i.e., correlate strongly enough) that those variables can be concatenated. In this case, the variables used as input were time points. PCA was used to identify distinct windows of time in the ERP averages (hereafter, temporal factors or "virtual windows") during which similar voltage variance was registered across consecutive sampling points.

The averaged ERP data for the Semantically-Related and Unrelated conditions[1] were combined into a data matrix comprised of 401 columns (one column for each of the sampling points spanning from 0 to 800 ms) and 3,416 rows (the averaged ERP voltages for 28 participants, at each of the 61 active recording electrodes excluding the Left Mastoid reference, in each of the two conditions). This matrix was used as input to a covariance-based, temporal PCA. In order to determine how many dominant-variance components were extracted by the PCA, we used Rule M (Preisendorfer & Mobley, 1998). Rule M estimates how many components extracted from a real data set account for more variance than corresponding components extracted from a data set of normally-distributed, randomly-sampled noise having the same dimensions as the real data set. All components meeting this criterion for each PCA were retained and rotated to simple structure using Promax (Hendrickson & White, 1964) with Kaiser normalization and k=2

[1]The ERP data related to the Phonological aspect of the task (Phonologically-Related versus P-Unrelated) were analyzed separately using this same procedure.

(Richman, 1986; Tataryn, Wood, & Gorsuch, 1999). All PCA procedures were completed using the Matlab-based PCA Toolbox (Dien, 2005).

Each of the temporal factors extracted from the data set describes a virtual time window during which a distinct pattern of voltage was active. The time-course of each temporal factor is described by a set of factor loadings. The ERP variance captured by each temporal factor is described by a set of factor scores. The factor scores summarize the average amplitude of ERP activity, for each participant, at each electrode, in each condition, within the virtual time window formed by each temporal factor.

As noted above, with traditional amplitude measurements the average amplitude of the ERP waveforms is determined within specified time windows for each participant, in each condition, at specific electrodes, and then analyzed statistically. Here, factor scores associated with specific temporal factors (i.e., those whose time-course was consistent with the N400 ERP component) - rather than average windowed amplitude measures - were computed for each participant and then analyzed statistically, in two steps. First, factor scores were computed for each participant, in each condition, at three midline electrodes (Fz, Cz, Pz) and submitted to a repeated-measures ANOVA with Condition entered as a within-subjects factor having two levels (Related, Unrelated), electrode entered as a within-subjects factor having three levels (Fz, Cz, Pz), and Group entered as a between-subjects factor having two levels (AWS, AWNS). This level of analysis gave a first-pass assessment of whether Condition and Group effects were present in the ERP voltages during the specified time window. Next, a more extensive topographic analysis of the ERP activity in that same "virtual window" was conducted. As noted above, for each temporal factor a set of factor scores is generated describing the average voltage activity during a specific time window for each participant, in each condition, at each electrode. Factor scores computed for each condition, at specific electrodes, were grouped and averaged. The electrodes were grouped into eight regions of interest (see Figure 6, based on the recommendations of Dien & Santuzzi, 2005): Left Anterior Inferior (FP1,F7,F5,FT7,FC5), Left Anterior Superior (AF3,F3,F1,FC3,FC1), Left Posterior Inferior (TP7,CP5,P7,P5,O1), Left Posterior Superior (CP3,CP1,P3,P1,PO3), Right Anterior Inferior (FP2,F8,F6,FT8,FC6), Right Anterior Superior (AF4,F4,F2,FC4,FC2), Right Posterior Inferior (TP8,CP6,P8,P6,O2), and Right Posterior Superior (CP4,CP2,P4,P2,PO4). The averaged factor scores were then submitted to a repeated-measures ANOVA with Hemisphere entered as a within-subjects factor having two levels (Left, Right), Dorsality entered as a within-subjects factor having two levels (Inferior, Superior), Anteriority entered as a within-subjects factor having two levels (Anterior, Posterior), Condition entered as a within-subjects factor having two levels (Related, Unrelated), and Group entered as a between-subjects factor having two levels (AWS, AWNS). For any test violating the assumption of sphericity we report p-values based on adjusted degrees of freedom (Greenhouse & Geiser, 1959) along with original F-values. When a statistically significant main effect or interaction was detected, Bonferroni-protected pair-wise comparisons were made.

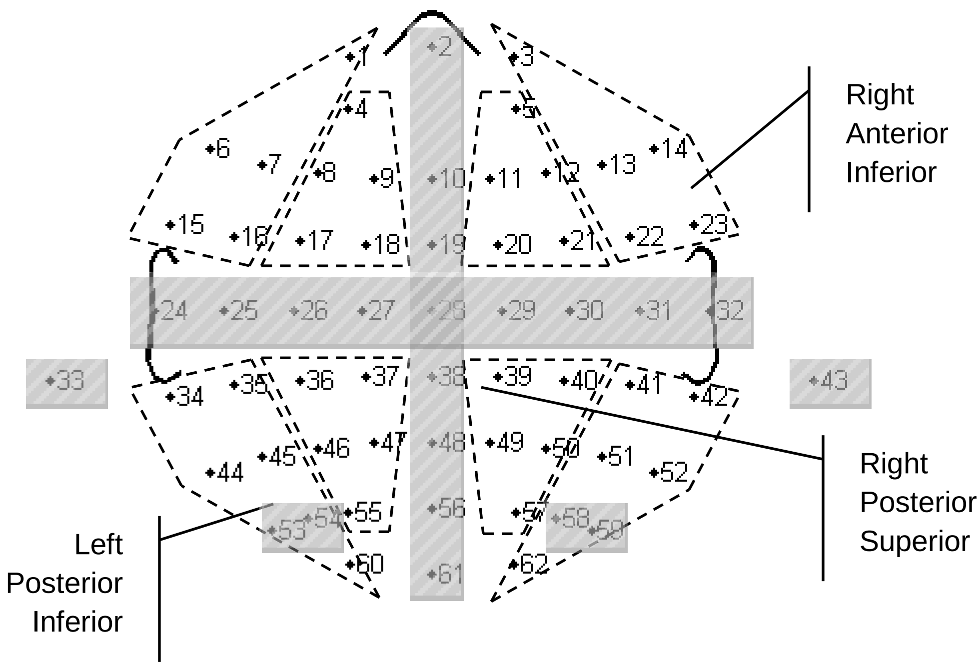

Figure 6. Eight topographic regions of interest (ROIs), bounded by hashed lines, each defined by five electrodes (each indicated here by a number; standard electrode names are given in the text). Electrodes obscured by shading were not included in any ROI.

RESULTS

Behavioral Data

Naming accuracy*.* Table 1 summarizes the naming accuracy for each group in each condition. Inspection of Table 1 suggests little difference in naming accuracy between Groups or between Conditions. Repeated-measures ANOVA confirmed that there was no main effect of Condition (F[4,104=0.847,p=0.498) and no interaction of Condition and Group (F[4,104]=0.600,p=0.637). However, the main effect of Group was statistically significant (F[1,26]=4.59,p=0.042), with the AWS less accurate than the AWNS; although the size of this effect was small (partial eta-squared = 0.15).

Group		Total Ave	Filler	S-Related	S- Unrelated	P-Related	P-Unrelated
AWS	**Mean**	96.94	96.03	96.03	95.50	95.77	96.56
	SD	2.84	3.26	2.70	3.75	3.66	2.98
AWNS	**Mean**	96.94	98.01	97.62	98.15	97.62	98.15
	SD	2.84	1.49	1.69	2.41	2.23	2.06

Table 1. Mean percent correct for naming accuracy by group and condition.

Word check accuracy. Table 2 summarizes the word check accuracy for each group in each probe word condition. Repeated-measures ANOVA detected no main effect of Condition (F[3,78]=.316,p=.761), no main effect of Group (F[1,26]=2.795,p=.107), and no interaction of Condition and Group (F[3,78]=2.189,p=.114).

Group		Total Ave	Semantically Related	Semantically Unrelated	Phonologically Related	Phonologically Unrelated
AWS	**Mean**	98.71	98.94	97.62	97.62	98.68
	SD	2.21	2.02	4.38	2.85	1.53
AWNS	**Mean**	98.71	98.68	99.47	99.47	99.20
	SD	2.21	1.98	0.87	0.87	1.20

Table 2. Mean percent correct for word check accuracy by group and condition.

Visual Inspection and PCA of ERP Data for Semantic Picture-Word Priming

Grand average waveforms for both Groups are shown in Figure 7 at three midline electrodes (Fz, Cz, Pz) for the Semantically-Related and Unrelated. Visual inspection reveals that both trial types elicited a similar sequence of ERP activity. Most notable is the later, negative-going activity indicative of N400. N400 priming (i.e., a reduction in amplitude for Semantically-Related versus Unrelated) is apparent for both groups, spanning the time window between ~500-800 ms.

Temporal PCA revealed 13 virtual time windows of interest. The time-course of each virtual window is defined by a set of factor loadings, shown in Figure 8. Consecutive time points with high loadings define the duration of time covered by each temporal factor (time points with relatively small loadings are ignored). The time point with the largest loading indicates the peak latency of each temporal factor. Each factor will, hereafter, be referred to according to its peak latency (e.g., T112, T234, etc). As shown in Figure 8, the temporal factors had peak latencies spanning from 14 ms to 786 ms. As noted above, N400 priming seemed to cover the 500-800 ms time window in the grand averaged data. We analyzed the ERP variance that was active during T434, T504, T760 and T786, respectively. However, statistically significant effects were only

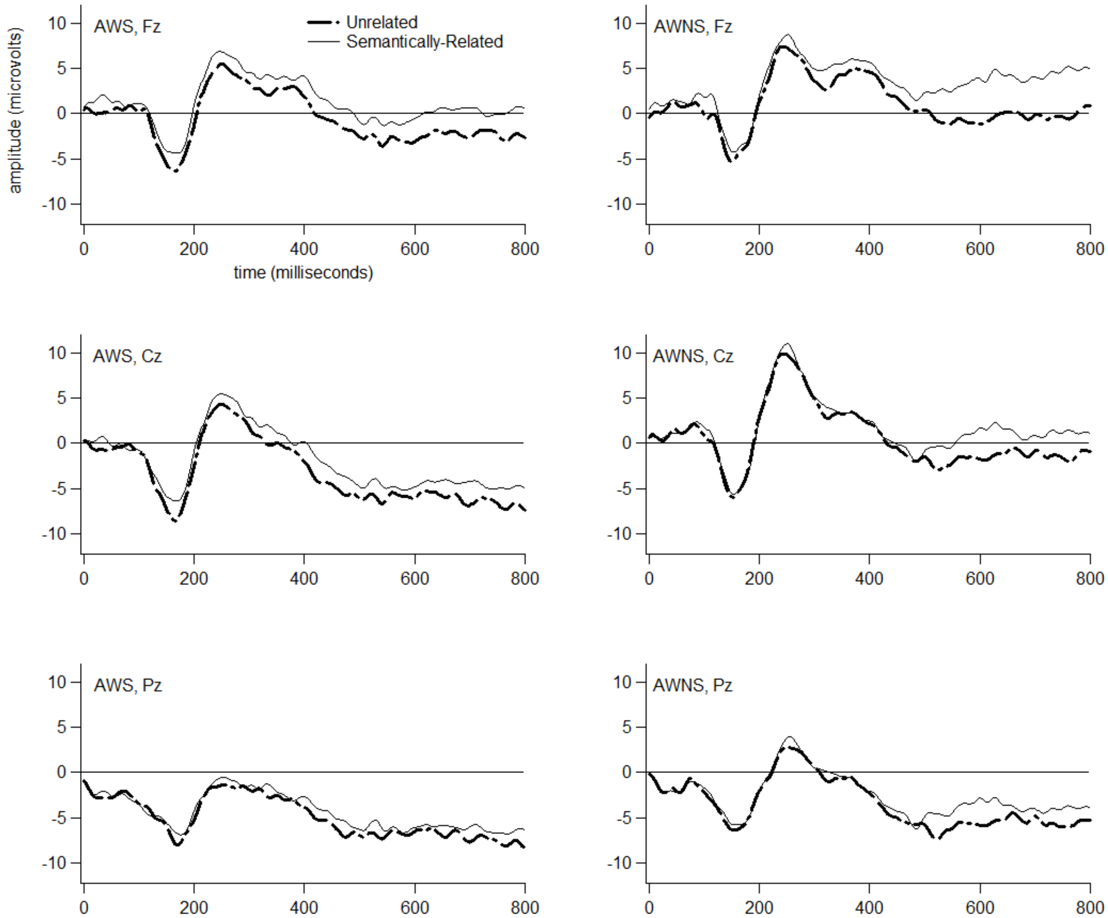

Figure 7. Grand average ERP waveforms for Semantically-Unrelated Probe Words, and Semantically-Related Probe Words.

observed during the T760 time window, as reported next. Neither the analysis at midline electrodes, nor the topographic analysis uncovered significant effects within the other three time windows.

Analysis of T760 activity at midline electrodes. T760 factor scores (summarizing the ERP variance during the T760 time window), generated for each group in each condition at electrodes Fz, Cz and Pz, were submitted to repeated-measures ANOVA as described previously. A statistically significant main effect of Electrode was detected ($F[2,52]=21.21$, $p=0.000$), as was a main effect of Condition ($F[1,26]=4.47$, $p=0.044$) and an interaction of Condition and Electrode ($F[2,52]=13.77$, $p=0.000$).

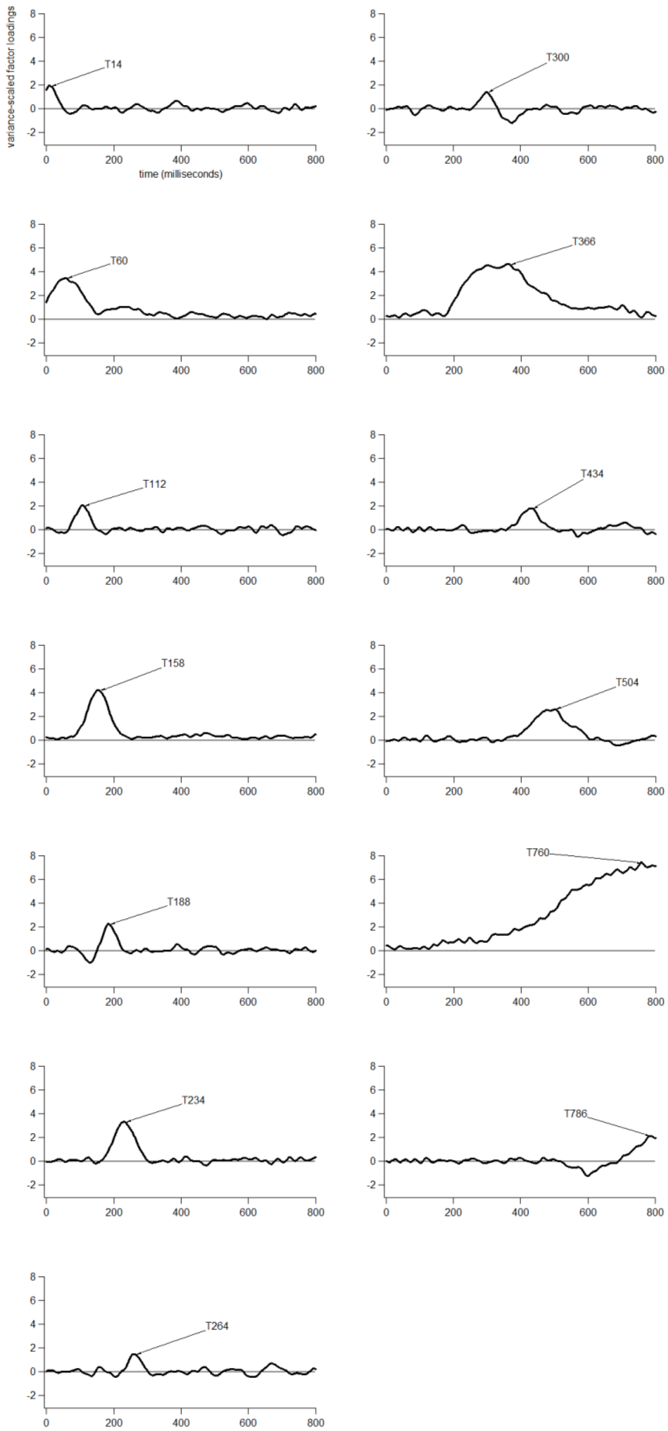

Figure 8. Factor loadings indicating the time-course of 13 temporal factors derived from the Semantically-Related/Unrelated data set. Peak latencies are given with each label (e.g., T264 = a peak latency of 264 ms).

Bonferroni-protected pair-wise comparisons of the Semantically-Related versus Unrelated scores at each electrode revealed that, at electrode site Fz only, Semantically-Related scores were attenuated relative to Unrelated (p=0.006). A marginally-significant Condition effect was also detected at electrode site Cz, with Semantically-Related scores attenuated relative to Unrelated scores (p=0.081). These effects are shown in Figure 9. No main effect of Group or interactions involving Group were detected.

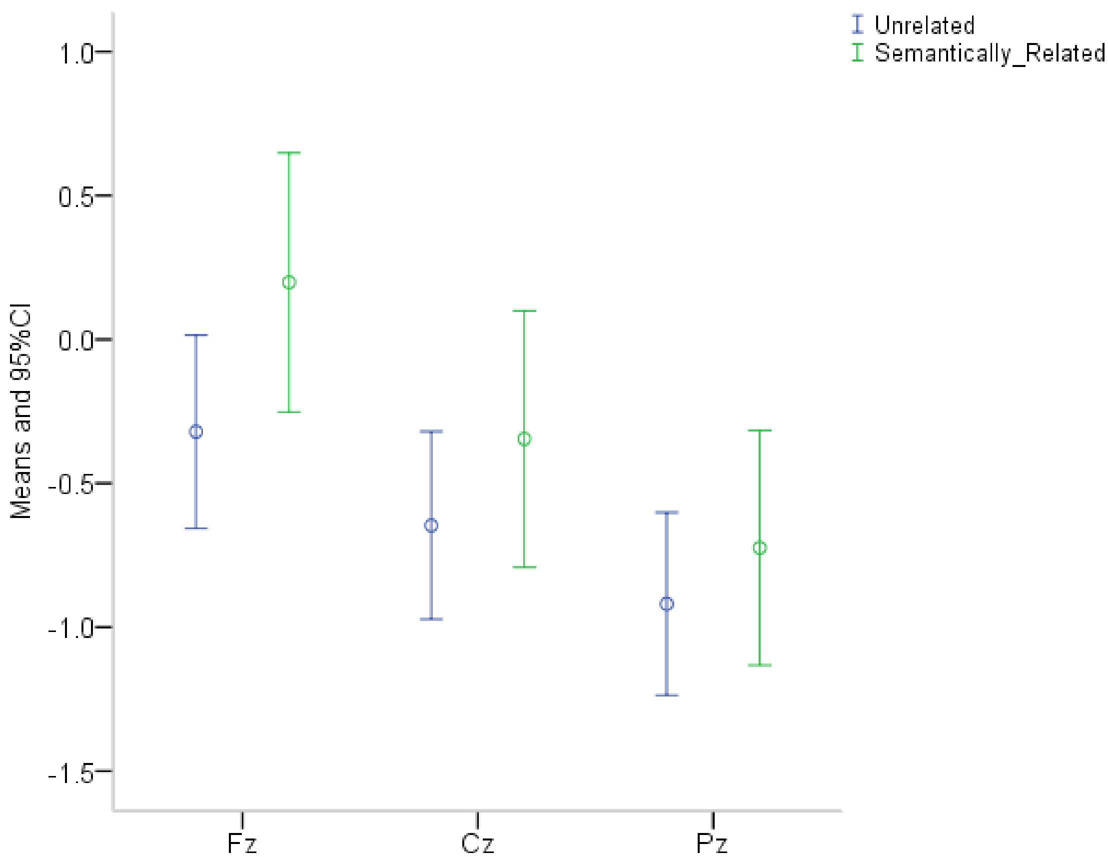

Figure 9. T800 factor scores (means and 95% confidence intervals) shown at each Electrode for each Condition.

Topographic analysis of T760 activity. T760 factor scores (summarizing the ERP variance during the T760 time window), generated for each group in each condition at each of the eight regions of interest, were submitted to repeated-measures ANOVA as described above. A statistically-significant five-way interaction of Group, Condition, Laterality, Dorsality, and Anteriority was detected ($F[1,26]=6.75, p=0.015$). Bonferroni-protected pair-wise comparisons revealed that, for the AWNS group only, the scores for Semantically-Related were attenuated versus Unrelated at the Left Anterior Superior region ($p=0.024$), at the Right Anterior Superior region ($p=0.021$), and (marginally) Right Anterior Inferior region ($p=0.057$). No statistically-significant Condition effects were detected for the AWS group at any of the eight topographic regions of interest. Thus, although we detected a Condition effect at Fz that generalized to both groups (see above section), we suspect that that effect was driven by the AWNS group, who evidenced a robust Anterior Semantic N400 priming effect.

Visual Inspection and PCA of ERP Data for Phonological Picture-Word Priming

Grand average waveforms for both Groups are shown in Figure 10 at three midline electrodes (Fz, Cz, Pz) for Phonologically-Related and P-Unrelated. Visual inspection reveals that both trial types elicited an overall similar sequence of ERP activity, with one apparent Group difference in the morphology of the later, negative-going activity indicative of N400. Whereas a typical N400 priming effect (i.e., a reduction in amplitude for Phonologically-Related versus P-Unrelated) is apparent for the AWNS group in the time window between ~500-650 ms, a reversal of this effect is apparent for the AWS group in the time window between ~400-500 ms.

Temporal PCA revealed 12 virtual time windows of interest. The time-course of each virtual window, given by its factor loadings, is shown in Figure 11. Each factor will, hereafter, be referred to according to its peak latency (e.g., T156, T202, etc). As shown in Figure 11, the temporal factors had peak latencies spanning from 10 ms to 800 ms.

As noted above, a typical N400 priming effect seemed evidence for the AWNS group in the time window between ~500-650 ms. We analyzed the ERP variance that was active during T448, T532 and T800, respectively. Statistically significant effects were detected in the first two of these three time windows. Neither the analysis at midline electrodes, nor the topographic analysis uncovered significant effects within the T800 time window.

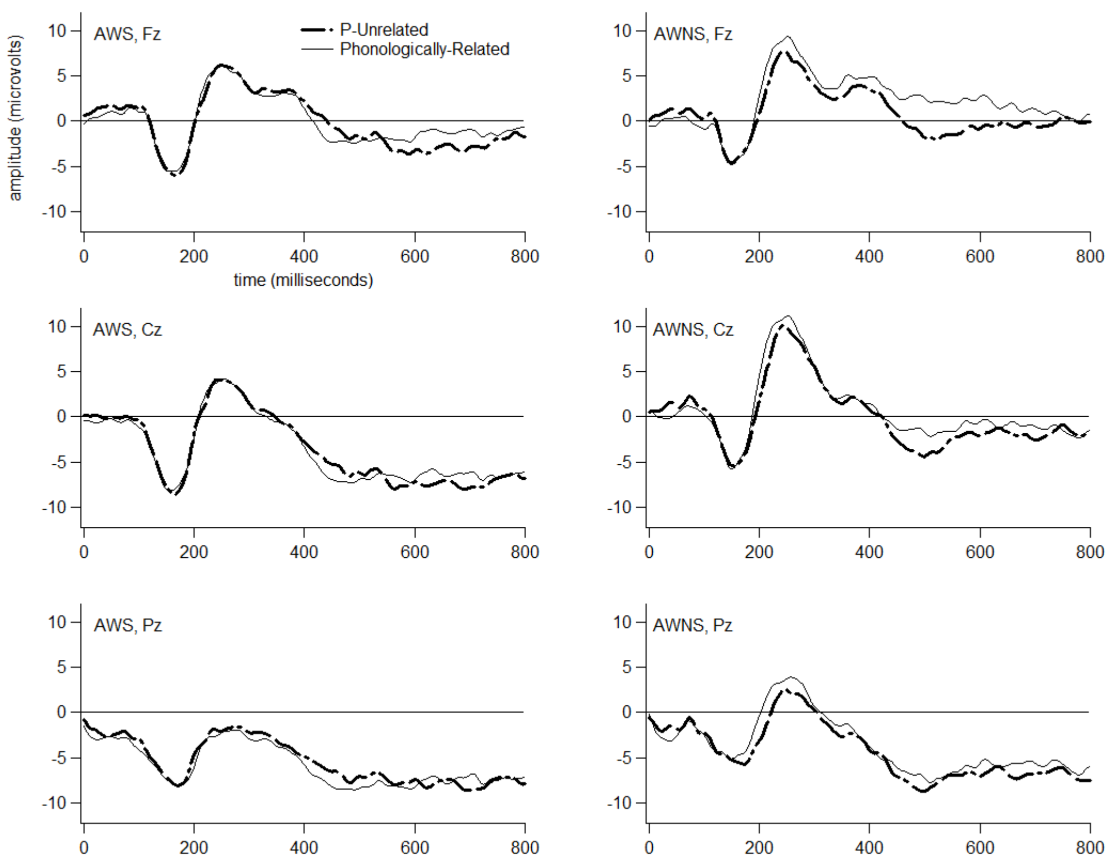

Figure 10. Grand average ERP waveforms for Phonologically-Unrelated Probe Words, and Phonologically-Related Probe Words.

Analysis of T448 activity at midline electrodes. T448 factor scores (summarizing the ERP variance during the T448 time window), generated for each group in each condition at electrodes Fz, Cz and Pz, were submitted to repeated-measures ANOVA. A statistically significant main effect of Electrode was detected ($F[2,52]=41.98$, $p=0.000$), as was an interaction of Electrode and Group ($F[2,52]=7.08$, $p=0.007$), an interaction of Condition and Group ($F[1,26]=6.91$, $p=0.014$) and - *most notably* - an interaction of Condition, Electrode and Group ($F[2,52]=4.92$, $p=0.016$). Bonferroni-protected comparisons of the two Conditions, at each electrode for each group, revealed that - for the AWS - Phonologically-Related scores had a larger negative-going amplitude than P-Unrelated scores at Fz ($p=0.013$), Cz ($p=0.043$) and Pz ($p=0.048$); while - for the AWNS - the P-Unrelated scores had a larger negative-going amplitude than Phonologically-Related scores at Fz only ($p=0.035$). These effects are shown in Figure 12.

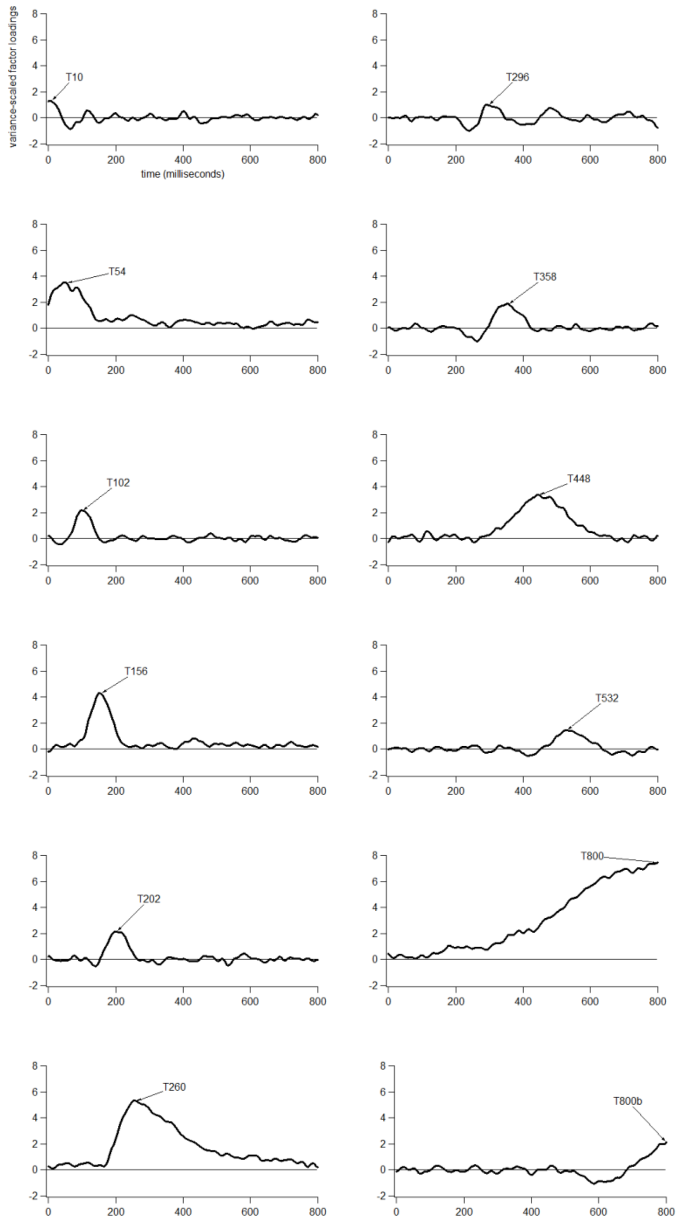

Figure 11. Factor loadings indicating the time-course of 12 temporal factors derived from the Phonologically-Related/P-Unrelated data set. Peak latencies are given with each label (e.g., T202 = a peak latency of 202 ms).

Topographic analysis of T448 activity. T448 factor scores (summarizing the ERP variance during the T448 time window), generated for each group in each condition at each of the eight regions of interest, were submitted to repeated-measures ANOVA. A statistically-significant interaction of Group and Condition was detected ($F[1,26]=6.17$, $p=0.02$), with Bonferroni-protected pair-wise comparisons revealing that Phonologically-Related had a larger negative-going amplitude than P-Unrelated for the AWS ($p=0.013$); an effect not detected for the AWNS ($p=0.409$).

Analysis of T532 activity at midline electrodes. T532 factor scores (summarizing the ERP variance during the T532 time window), generated for each group in each condition at electrodes Fz, Cz and Pz, were submitted to repeated-measures ANOVA. A statistically significant interaction of Electrode and Condition was detected ($F[2,52]=8.50$, $p=0.002$), as was an interaction of Electrode, Condition and Group ($F[2,52]=5.14$, $p=0.018$). Bonferroni-protected pairwise comparisons of the Phonologically-Related versus Unrelated Scores for each group at each electrode revealed that - for the AWNS group only - Phonologically-Related scores had an attenuated negative-going amplitude relative to Unrelated at electrode Fz ($p=0.012$) (see Figure 13).

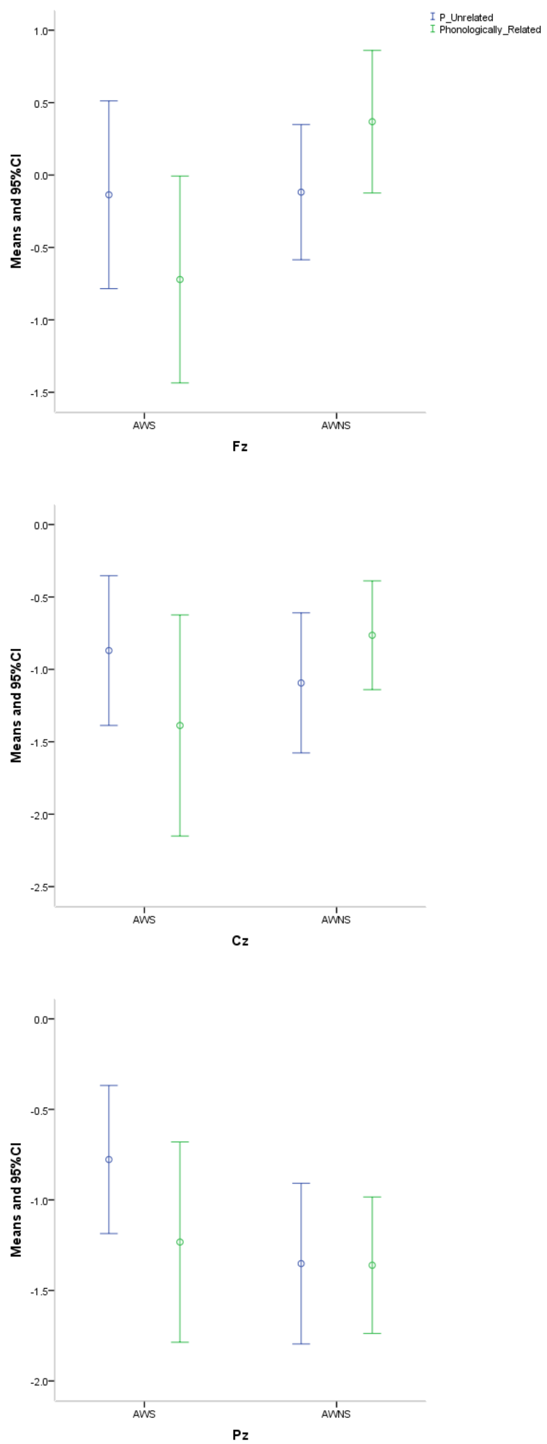

Figure 12. T448 factor scores (means and 95% confidence intervals) shown for each Group (AWS on the left, AWNS on the right), in each Condition (blue = P-Unrelated, green = Phonologically-Related), at each Electrode (Fz, Cz and then Pz, respectively).

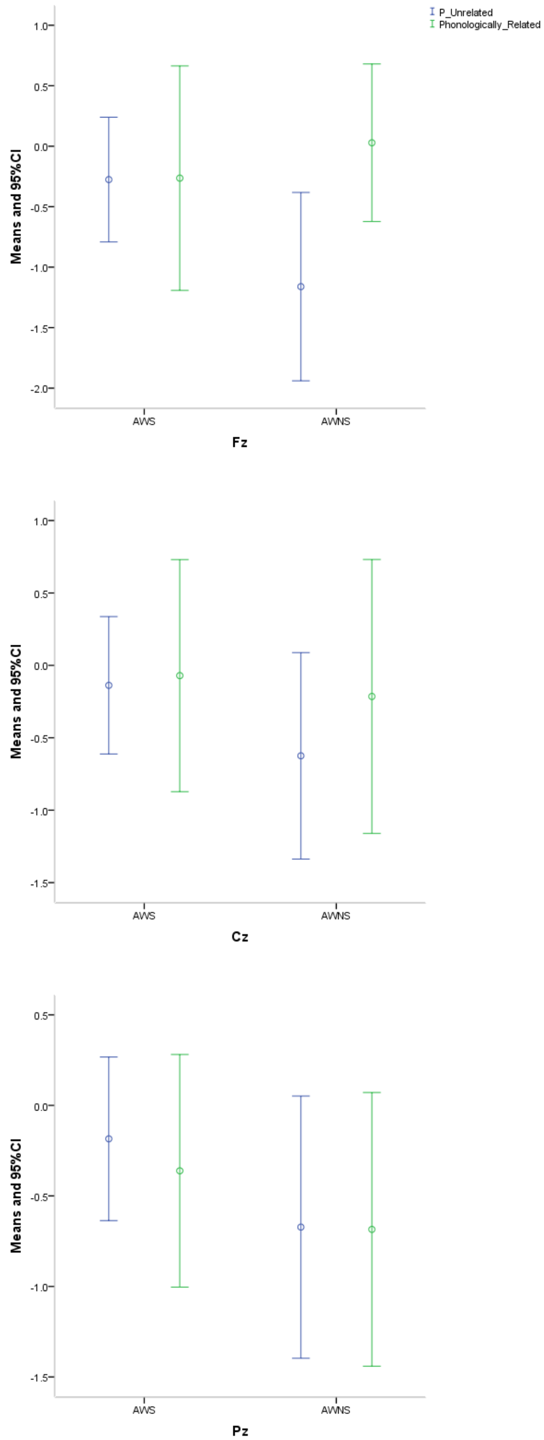

Figure 13. T532 factor scores (means and 95% confidence intervals) shown for each Group (AWS on the left, AWNS on the right), in each Condition (blue = P-Unrelated, green = Phonologically-Related), at each Electrode (Fz, Cz and then Pz, respectively).

Topographic analysis of T532 activity. T532 factor scores (summarizing the ERP variance during the T532 time window), generated for each group in each condition at each of the eight regions of interest, were submitted to repeated-measures ANOVA. A statistically-significant five-way interaction of Group, Condition, Laterality, Dorsality, and Anteriority was detected ($F[1,26]=4.27, p=0.049$). Bonferroni-protected pairwise comparisons revealed that - for the AWNS group only - Phonologically-Related scores had an attenuated negative-going amplitude relative to P-Unrelated at the Left Anterior Inferior region ($p=0.039$), at the Left Anterior Superior region ($p=0.034$), and (marginally) at the Right Anterior Superior region ($p=0.054$). No statistically-significant Condition effects were detected for the AWS group at any of the eight topographic regions of interest.

Discussion

The purpose of this study was to assess semantic and phonological activation spreading in AWS on the path to picture naming. Fourteen AWS and 14 AWNS participated in a picture-word priming task. A picture-to-be-named presented on each trial was sometimes followed by an auditory probe word. Some of the probe words were Semantically-Related (but phonologically-unrelated) to the target labels of their preceding pictures. Those same probe words also appeared following pictures with labels that were semantically- and phonologically-unrelated (Unrelated). ERPs were recorded to both of these probe word categories and compared (Semantically-Related versus Unrelated ERP comparison). Still other probe words presented during the task were Phonologically-Related (but semantically-unrelated) to the target labels of their preceding pictures. Those same probe words also appeared following pictures with labels that were semantically- and phonologically-unrelated (P-Unrelated). In a separate analysis, ERPs recorded to these probe word categories were compared (Phonologically-Related versus P-Unrelated ERP comparison). Our aim was to determine whether predictable N400 amplitude fluctuations were present between specific conditions, and similar in magnitude, for each group. Specifically, we examined, and compared between-groups, the magnitude of N400 amplitude shift in the Related versus Unrelated conditions and, separately, in the Phonologically-Related versus P-Unrelated conditions.

Semantic N400 Priming in the AWNS: Evidence of Semantic Activation Spreading

As predicted based on previous research, the AWNS evidenced N400 priming (i.e., attenuation of the N400 component) for Semantically-Related versus Unrelated probe words. This effect was robust for the AWNS at the frontal midline electrode (Fz), and at three larger regions of the scalp, during a relatively late time window peaking in amplitude at ~760 ms. The three scalp regions at which Semantic N400 priming was detected in our AWNS group were the frontal anterior superior regions (left and right) and the right anterior inferior region.

Although this effect was comparable to those found in past studies using a similar research design, the topography of the Semantic N400 priming effect for our AWNS was notably focal. Jescheniak et al. (2002) elicited two qualitatively different Semantic N400 priming effects. One was detected between ~400-800 ms, and was broadly distributed over the scalp when participants were required to name the pictures while also attending to the auditory probe words (see Jescheniak et al., 2002, Experiment 1). Their other Semantic N400 priming effect was also detected between ~400-800 ms, but had a more anterior distribution, when participants were required to make a size judgment about the pictures (instead of naming them) while also attending to the auditory probe words (see Jescheniak et al., 2002, Experiment 2). Jescheniak et al. (2002) surmised that the broadly-distributed Semantic N400 priming effect seen in their first experiment reflected lexical-semantic priming, while the anteriorly-focused Semantic N400 priming effect seen in their second experiment probably reflected conceptual priming since access to the picture names was not required in order to make the size judgments.

Since the Semantic N400 priming effect seen in the current study for our AWNS group was also frontally-focused, one speculation is that those participants processed the probe words with less depth at a lexical-semantic level. One reason for this might have been the intense focus of our task on phonological rehearsal of the auditory probe words. Recall that, during the word check, participants were presented with a written word that was either the probe or a word that was phonologically very similar, and had to indicate whether the written word matched the probe word for that trial. We speculate that the word check design (i.e., the phonological similarity between matches and mismatches) caused participants to attend to and process the auditory probe words at such a fine-grained level, phonologically, that they processed the probe words with less depth semantically. Thus, although Semantic N400 priming was observed for our AWNS, it was not as widely-distributed across the scalp as it was, e.g., in Jescheniak et al. (2002, Experiment 1). Maxfield et al. (in press) also reported Semantic N400 priming effects that were more broadly distributed for their AWNS group than that seen here using the same equipment, similar picture and stimuli (but with a different design), and similar participants to the current study. Thus, it would appear that the differences between this study and Jescheniak, et al. (2002) are due to differences in the experimental design.

Non-Robust Semantic N400 Priming in the AWS: Evidence of Weak Semantic Activation Spreading

As discussed in the paragraph above, a Condition by Electrode interaction was detected in the midline analysis of Semantically-Related versus Unrelated ERPs: At electrode Fz only, during a time window peaking at ~760 ms, N400 attenuated significantly for Semantically-Related versus Unrelated, an effect that generalized statistically to both groups. Although the three way interaction (i.e., Group X Condition X Electrode) was not significant, we suspected that the Semantic N400 priming effect at Fz was driven primarily by our AWNS group. Bonferroni-corrected pairwise comparisons of Semantically-Related versus Unrelated at electrode Fz separately for each group (not reported above in the Results), confirmed that Semantic N400 priming was statistically significant for the AWNS (as discussed in the section above) but not for the AWS. What is possible is that some but not all of the AWS showed a Semantic N400 priming effect at Fz, such that the three-way interaction of Condition, Group and Electrode was not significant. The topographic analysis (including Regions as factors) corroborates this interpretation: A statistically significant Semantic N400 priming effect was not detected at any of the eight non-midline scalp regions in the AWS group, i.e., whatever Semantic N400 priming effect was present for our AWS group at electrode Fz did not volume-conduct to other scalp regions as it did for our AWNS group.

We can envision at least two interpretations for the non-robust Semantic N400 priming effect in our AWS group. One is that semantic activation spreading was non-robust in at least some of our AWS. That is, on the path to picture naming in at least some of our AWS, activation may have spread only weakly to a cohort of words in the mental lexicon semantically-related to the target picture label. Weaker semantic activation spreading, in turn, may indicate that the semantic network structure of the mental lexicon is deficient in AWS. As reviewed in the Introduction, several other published findings support this conclusion. Our results converge with theirs to suggest that at least some AWS have some level of difficulty organizing and activating words according to their semantic features.

An alternate interpretation of the non-robust Semantic N400 priming seen in our AWS group is that it reflects a phonological rehearsal deficit. As discussed above, attentional control was heavily focused on phonological rehearsal of the auditory probe words during our task. Participants were to remember both the picture label for naming, and the auditory probe word for the word check. Above we argued that the design of the word check required participants to pay particularly close attention to phonological features of the auditory probe words. This seems to have resulted in a relatively focal albeit still robust Semantic N400 priming effect in our AWNS. Perhaps, in our AWS, phonological rehearsal of the probe words was so demanding that they processed the probes with little to no depth semantically - resulting in a very weak Semantic N400 priming effect for this group. Bajaj (2007) reviews evidence that phonological working memory is limited in AWS, an effect that may have caused them to divert resources away from semantic processing of the auditory probe words. A way to test this hypothesis in a future experiment would be to replicate the task but with a phonologically less demanding word check. If a more robust Semantic N400 priming effect were seen in the AWS, the conclusion that phonological rehearsal of the probe words was particularly demanding in the current task might garner additional support. If, on the other hand, Semantic N400 priming continued to be weak or absent, the conclusion that semantic activation spreading was weak in at least some of our AWS in the current task might be more likely.

Phonological N400 Priming in the AWNS: Evidence of Phonological Activation Spreading

The AWNS also evidenced N400 attenuation for Phonologically-Related versus Unrelated probe words; an effect that was most robust during a time window peaking at ~532 ms after word onset at electrode Fz and at three anterior scalp regions. This effect is consistent with the Phonological N400 priming effect reported for AWNS in Jescheniak et al. (2002, see Experiment 1). It is noteworthy however that, in their experiment, Phonological N400 priming had slightly different scalp topography. Whereas in our AWNS group Phonological N400 priming was evidenced at the Left Anterior Inferior region, Left Anterior Superior region and - weakly - at the Right Anterior Superior scalp region; Jescheniak et al. (2002) reported a Phonological N400 priming effect that was not robust at Left Anterior and Left Central regions. We are unsure of what may have driven this topographical difference, or of what it may mean functionally.

Reverse Phonological N400 Priming in the AWS

In contrast, the AWS evidenced a reversal of the Phonological N400 priming effect, i.e., N400 amplitude was significantly larger for Phonologically- Related versus P-Unrelated probe words. This effect was detected during a time window peaking at ~448 ms after word onset - roughly 50 ms earlier than the typical Phonological N400 priming effect seen in our AWNS group. The reversed Phonological N400 priming effect for the AWS also had a widespread topographical distribution, i.e., it was detected at all three midline electrodes and generalized to all eight scalp regions analyzed topographically.

A search of the literature reveals no reports of a reversal of the N400 effect involving phonological priming in the typical population. However, a small body of work has reported that there are conditions in which semantic priming can result in a reversal of the N400 effect. As reviewed in Maxfield et al. (in press), one study demonstrated that a weakly-activated prime word can trigger a "center-surround inhibition" process, whereby semantically-related neighbors of the weakly-activated prime word are inhibited so that the prime word can be accessed more fully (Bertmeitinger, Frings, & Wentura, 2008). In that study, a special masked priming method was used to ensure that the prime words would become only weakly activated. However, there are potentially other reasons that prime words - which, for us, were the self-generated picture labels - might become only weakly activated during a lexical task. One possibility is that target picture labels became only weakly activated in our AWS group on the path to picture naming, causing them to inhibit other neighboring words so that the target picture labels could be fully accessed. As reviewed in the Introduction, there is some evidence to suggest that AWS have difficulty activating target words in the mental lexicon. A problem with this interpretation is that it is not clear why Phonologically-related but not Semantically-related neighbors of the target picture labels should be inhibited if AWS have difficulty activating target words in the mental lexicon, i.e., both types of priming should have produced a reverse N400 priming effect in our task. We note that a reverse Semantic N400 priming effect was seen for AWS in Maxfield et al. (in press) when less attention was paid to the auditory probe words. Thus, it **is** possible to elicit reverse Semantic N400 priming in AWS. However, we have seen here that when greater attention is paid to the auditory probe words (and, specifically, to their phonological structure), N400 reversal is seen in AWS for Phonological priming instead of for Semantic priming. Thus, attentional control may be a factor in generating this effect in AWS on the path to picture naming.

This leads to a second interpretation of the reversal Phonological N400 priming effect seen in the current study for our AWS group; namely, that it reflects an attentional control deficit. As noted above, there is ample evidence to suggest that AWS have phonological working memory deficits (see Bajaj, 2007). Maxfield et al. (in press) reviewed another study demonstrating that AWNS strategically inhibit words under certain monitoring conditions; which results in a reversal of the typical N400 priming effect (Mari-Beffa, Valdes, Cullen, Catena & Houghton, 2005). It is possible that, in our task, AWS may have relied on a similar, executive approach to ensure that processing resources were allocated in such a way as to maximize their performance given presumed phonological working memory limitations. Specifically, as AWS tried to hold each target picture label in mind, it seems plausible that - when a phonologically-related probe word was presented - they may have used a processing strategy of inhibiting activation of Phonologically-Related probe words so that these words would not interfere with the maintenance of the target picture labels in phonological working memory. We suspect that this effect is what resulted in larger-than-usual N400 amplitude on Phonologically-Related trials for our AWS. It is interesting to note that, in contrast to our results, Weber-Fox, Spencer, Spruill, and Smith (2004) found similar N400 priming effects to Phonologically-Related words in AWS versus AWNS in a task that did not require overt speech, suggesting that the phonological encoding system may operate similarly in AWS versus AWNS in the absence of speaking demands. Those authors did observe that phonological N400 priming was more lateralized to the right hemisphere of the scalp for their AWS group versus their AWNS group. Overall, though, our results in light of theirs lead us to speculate that the demand of managing delayed picture naming plus auditory probe word monitoring in our task is what contributed to the N400 enhancement we saw on Phonologically-Related trials for our AWS group.

Study Limitations

It is important to acknowledge a potential limitation of the research reported here. That is, our groups differed in the proportion of females versus males. Whereas the AWS group was comprised primarily of male participants, a large proportion of female participants constituted the AWNS group. Gender has been shown to influence the topographic distribution of the N400 ERP component in the following ways. For Semantic N400 priming, Van Dyke et al. (2010) found that women demonstrate evidence of bilateral language processing; whereas men demonstrate language processing that is left hemisphere dominant. For Phonological N400 priming, Lindell and Lum (2008) found that females had consistently faster response times than males in both experiments of word/nonword judgment and rhyming tasks. Although that task did not involve neural imaging, those authors surmised that the faster response time for females resulted because greater neural resources were devoted to processing the stimulus items; whereas the males were relying on solely the left hemisphere. Given these effects, one minor concern with our study was that the differences in scalp topography observed between groups for our N400 effects might have been driven by the different gender distributions in our two groups. However, our topographic results were not consistent with previously-reported gender differences. Still, at this time it is difficult to know how much gender might have been a factor in driving the topographic differences seen between groups in our study.

Clinical Implications

Since our AWS group demonstrated differences versus the AWNS control group, this suggests that AWS do not process language typically; particularly when language tasks require greater attentional control. Perhaps, as part of therapy, attentional training should be employed focusing on split attention, attention switching and sustained attention during word processing. For example, Weiss (2004b) discusses that naturalistic social speaking tasks should be included in therapy for stuttering in order to enhance vocabulary, pragmatics, discourse, conversation, and generalization of fluency-enhancing techniques. According to Weiss (2004b), a good starting point is with easy situations (e.g., monologueing and story retells) and working up to a more difficult situation (e.g., multipartner conversation). In multipartner conversation, therapy would focus on pragmatics, topic maintenance, topic flow, pragmatics, comprehension of the conversation, and concomitant attention skills. If it can be confirmed that attention is the influencing factor with our AWS, another potentially useful approach would be to train attention directly using a software-based drill-like therapy technique. For example, the Attention Process Training (APT-I, Sohlberg & Mateer, 2001) from Aphasia therapy could be used.

Lexical structuring tasks such as word games, chaining, and semantic relation games might also help AWS improve the organization of their mental lexicons in such a way that access to words can be streamlined even when attentional demands are high. Weiss (2004a) argues that a language-based approach should be employed in stuttering therapy in order to enhance the linguistic skills of those who stutter on a continuum of increasing complexity, which has been shown to improve fluency generalization.

Due to the variability in results we believe we elicited in our data from the AWS, we considered that further subtyping of AWS should be explored. Yairi (2007) has already considered this and takes many life factors into consideration in subtyping such as personality, gender, genetics, severity, handedness, etc. We believe that our findings further reinforce the need to further classify AWS as the classifications will have an impact on therapeutic approaches.

Summary and Conclusions

We investigated semantic and phonological picture-word processing in AWS, hypothesizing that if the mental lexicon is well-organized and operates efficiently then we would evoke typical N400 priming effects for this group. Our AWNS group exhibited a typical Semantic N400 priming effect while a non-robust Semantic N400 priming effect was seen in our AWS. We interpreted the latter as possibly resulting from one of at least two different phenomena. One was that semantic activation spreading was weak in our AWS group secondary to a deficit in the semantic network structure of the mental lexicon in AWS; rendering weak Semantic N400 priming. A second interpretation was that the non-robust Semantic N400 priming effect reflected a phonological rehearsal deficit in the AWS. Both interpretations are tentative, and others are certainly possible. Above we proposed a future experiment for testing each hypothesis.

Our AWNS group also exhibited a typical Phonological N400 priming effect. In contrast, our AWS group presented with a reverse Phonological N400 priming effect. We interpreted this effect as likely reflecting an attentional control strategy used by the AWS to help them hold in working memory both the picture label and the auditory probe word on each trial - a task that might be particularly difficult in AWS, who have been shown to have phonological working memory deficits.

Our findings - considered alongside those of Maxfield et al. (in press) - suggest that AWS may process language differently than AWNS, particularly when the language task requires good attentional control. Whereas the AWNS tested in each experiment showed expected N400 priming effects, the AWS evidenced reverse N400 priming in one condition of each task. When phonological processing was emphasized, Phonological priming elicited a reversal of the N400 priming effect. When phonological processing was de-emphasized, Semantic priming elicited a reversal of the N400 priming effect. The extent to which these effects are artifacts of the task design remains to be seen. That is, it is unclear whether, during continuous speech, AWS - when trying to attend to a particular aspect of their message - might adopt this same strategy of inhibiting the activation of related words. What is clear is that, as suggested in the Introduction, top-down (i.e., central executive) influences can affect linguistic processing on the path to picture naming, and AWS seem to adopt executive (attentional-control) strategies not seen in AWNS. Further research is needed to understand the extent to which AWS adopt such strategies in other speaking contexts, and how adopting such controls might affect other aspects of psycholinguistic planning or speech motor performance.

References

Bajaj, A. (2007). Working memory involvement in stuttering: exploring the evidence and research implications. *Journal of Fluency Disorders*, *32*(3), 218-238. doi: 10.1016/j.jfludis.2007.03.002.

Bell, A. J. & Sejnowski, T. J. (1995). An information maximization approach to blind separation and blind deconvolution. *Neural Computation, 7*, 1129-1159.

Bermeitinger, C., Frings, C., & Wentura, D. (2008). Reversing the N400: Event-related potentials of negative semantic priming effect. *Cognitive Neuroscience and Neuropsychology, 19*(15), 1479-1482. doi: 10.1097/WNR.0b013e32830f4b0b.

Boberg, E., & Kully, D. (1994). Long-term results of an intensive treatment program for adults and adolescents who stutter. *Journal of Speech and Hearing Research, 37,* 1050-1059.

Bothe, A. K., Davidow, J. H., Bramlett, R. E., & Ingham, R. J. (2006). Stuttering treatment research 1970–2005: I. Systematic review incorporating trial quality assessment of behavioral, cognitive, and related approaches. *American Journal of Speech-Language Pathology, 15*, 321–341.

Burger, R. & Wijnen, F. (1999). Phonological encoding and word stress in stuttering and nonstuttering subjects. *Journal of Fluency Disorders, 24*, 91–106.

Caramazza, A., & Coltheart, M. (2006). Cognitive Neuropsychology twenty years on. *Cognitive Neuropsychology*, *23*(1), 3-12. doi:10.1080/02643290500443250.

Conture, E. G., (1996). Treatment efficacy: Stuttering. *Journal of Speech and Hearing Research, 39,* S18-S26. doi: 0022-4685/96/3905-OS18.

Craig, A. (1998). Relapse following treatment for stuttering: A critical review and correlative data. *Journal of Fluency Disorders, 23*, 1-30. doi: 0094-730X/98/$19.00.

Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders, 34*, 61-71. doi: 10.1016/j.fludis.2009.05.002.

Craig, A. R., & Hancock, K. (1995). Self-reported factors related to relapse following treatment for stuttering. *Australian Journal of Human Communication Disorders, 23,* 48-60.

Craig, A., Hancock, K., Tran, Y., Craig, M., and Peters, K. (2002). Epidemiology of stuttering in the community across the entire life span. *Journal of Speech, Language, and Hearing Research, 45*(6), 1097-1105. doi: 1092-4388/02/4506-1097.

Craig, A., & Tran, Y. (2005). The epidemiology of stuttering: The need for reliable estimates of prevalence and anxiety levels over the lifespan. *Advances in Speech-Language Pathology, 7*(1), 41-46. doi: 10.1080/14417040500055060.

Cutting, J. C. (1998). *Comprehension versus production.* In: Verschueren, J., Ostman, J., Blommaert, J., Bulcaen, C. (Eds.). The Handbook of Prgamatics*.* Philadelphia: John Benjamins Publishing Company.

Davis, S., Howell, P., Cooke, F. (2002). Sociodynamic relationships between children who stutter and their non-stuttering classmates. *Journal of Child Psychology and Psychiatry 43*, 939–947.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93(*3), 283-321.

Dell, G.S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*, 313-349.

Dell, G., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al (1991). *Psychological Review*, *98*(4), 604-614. doi:10.1037/0033-295X.98.4.604.

Dell, G.S. & O’Seaghdha, P.G. (1992). Stages of lexical access in language production. *Cognition, 42*, 287– 314.

Dien, J. (2005). PCA Toolbox [Computer software] (Version 1.93). Lawrence, KA. Available from http://www.people.ku.edu/~jdien/dowloads.html.

Dien, J. & Frishkoff, G. A. (2005). *Principal components analysis of event-related potential datasets.* In: Handy, T. (Ed.). Event-Related Potentials: A Methods Handbook. Cambridge: MIT Press.

Dien, J. & Santuzzi, A. M. (2005). Application of repeated measures ANOVA to high-density ERP datasets: A review and tutorial. In: Handy, T. C. (Ed.) *Event Related Potentials: A Methods Handbook,* 57-82.

Eldridge, K.A. & Felsenfeld, S. (1998). Differentiating mild and recovered stutterers from nonstutterers. *Journal of Fluency Disorders, 23*, 173–195.

Ferree, T.C., Luu, P., Russell, G.S. & Tucker, D.M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology, 112*, 536–544.

Fischler, I. (1990). *Comprehending language with event-related potentials*. In J. W. Rohrbaugh, R. Parasuraman & R. Johnson, Jr. (Eds), Event-Related Brain Potentials: Basic Issues and Applications, 165-177. New York: Oxford University Press.

Frazier, L. (1982). *Shared components of production and perception*. In: Caplan, D., Marshall, J. C. (Eds.). Neural Models of Language Processes. New York: Academic Press, 225-236.

Garrett, M.F. (1988). *Processes in language production*. In F. J. Newmeyer (Ed.), Linguistics: The Cambridge Survey (Volume 3, pp. 69-96). Cambridge, MA: Harvard University Press.

Glass, K., Frishkoff, G. A., Frank, R. M., Davey, C., Dien, J., Maloney, A. D. (2004). *A framework for evaluating ICA methods of artifact removal from multichannel EEG*. In: ICA, LNCS 3195. Berlin, Heidelberg: Springer-Verlagm, 1033-1040.

Glaser, W. R. (1992). Picture naming. *Cognition, 42*, 61-105.

Greenhouse, S. W. & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*, 95-112.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*(5), 350-365. doi:10.1016/j.jcomdis.2006.06.013.

Hagoort, P. & Kutas, M., (1995). Electrophysiological insights into language deficits. *Handbook on Neuropsychology, 10,* 105-134.

Hendrickson, A.E., & White, P.O. (1964). Promax: A quick method for rotation to oblique simple structure. *The British Journal of Statistical Psychology, 17*, 65-70.

Hennessey, N., Nang, C., & Beilby, J. (2008). Speeded verbal responding in adults who stutter: Are there deficits in linguistic encoding? *Journal of Fluency Disorders, 33*, 180-202.

Hockett, C. F. (1967). Where the tongue slips, there slip I. In: *To honor Roman Jakobson, Volume II (Janua Linguarum)*, 32, 910-936.

Hockett, C. F. (1973). Where the tongue slips, there slip I. In: Fromkin, B. A. (Ed.), *Speech errors as linguistic evidence* (93-119). The Hague: Mounton.

Howie, P. M., Tanner, S., & Andrews, G. (1981). Short-term and long-term outcome in an intensive treatment program for adult stutterers. *Journal of Speech and Hearing Disorders, 46*, 104-109. doi: 1092-5171/06/3301-0061.

Hugh-Jones, S., Smith, P. K. (1999). Self-reports of short- and long-term effects of bullying on children who stammer. *British Journal of Educational Psychololgy 69*, 141–158.

Jescheniak, J. D., Schriefers, H., Garrett, M. F., & Friederici, A. D. (2002). Exploring the activation of semantic and phonological codes during speech planning with event-related brain potentials. *Journal of Cognitive Neuroscience, 14*(6)*,* 951-964.

Kay, J., Coltheart, M. & Lesser, R. (1992). *Volume 3 of PALPA: Psycholinguistic assessments of language processing in aphasia.* Erlbaum.

Kempen, G. (2000). Could grammatical encoding and grammatical decoding be subserved by the same processing module? *Behavioral Brain Science, 23,* 38-39.

Kolk H. H. J. & Postma, A. (1997). *Stuttering as a covert-repair phenomenon*. In: Curlee, R. F., Siegel, G. (Eds.). Nature and Treatment of Stuttering: New Directions. Boston: Allyn & Bacon, 189-203.

Levelt, W. J. M., (1983). Monitoring and self-repair in speech. *Cognition, 14*, 41-104.

Levelt, W. J. M., (1989). *Speaking: From intention to articulation.* Cambridge, MA: The MIT Press.

Levelt, W., & Meyer, A. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, *12*(4), 433-452. doi:10.1080/095414400750050178.

Lindell, A. K. & Lum, J. A. G. (2008). Priming vs. rhyming: Orthographic and phonological representations in the left and right hemispheres. *Brain and Cognition 68*, 193-203. doi: 10.1016/j.bandc.2008.04.005.

Marí-Beffa, P., Valdés, B., Cullen, D., Catena, A., & Houghton, G. (2005). ERP analyses of task effects on semantic processing from words. *Cognitive Brain Research*, *23*(2-3), 293-305. doi:10.1016/j.cogbrainres.2004.10.016.

Maxfield, N., Huffman, J., Frisch, S., & Hinckley, J. (in press). Neural correlates of semantic activation spreading on the path to picture naming in adults who stutter. *Clinical Neurophysiology*. doi:10.1016/j.clinph.2010.03.026.

Miller, G.A. (1991). *The Science of Words*. New York: Scientific American Library.

Miller, L. & Lee, C. (1993). Construct validation of the Peabody Picture Vocabulary Test-Revised: A structural equation model of the acquisition order of words. *Psychological Assessment, 5*, 438-441.

Monsell, S. (1987). Nonvisual orthographic processing and the orthographic input lexicon. *Attention and performance 12: The psychology of reading*, 299-323. Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Retrieved from: http://www.usf.edu/FreeAssociation/.

Newman, R. & Bernstein Ratner, N. (2007). The role of selected lexical factors on confrontation naming accuracy, speed, and fluency in adults who do and do not stutter. *Journal of Speech, Language and Hearing Research, 50*, 196–213.

Nunez P. L., & Srinvasan, R. (2006). Electrical fields of the brain: The neurophysics of EEG (Appendices J1-J3). New York: Oxford University Press.

Nusbaum, H.C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No. 10*. Speech Research Laboratory, Bloomington IN. 357-376.

O'Halloran, R., & Larkins, B. (2008). The ICF Activities and Participation related to speech-language pathology. *International Journal of Speech-Language Pathology*, *10*(1-2), 18-26. doi: 10.1080/14417040701772620.

Otten L. J. & Rugg, M. D. (2005). *Interpreting event-related brain potentials*. In Handy, T. C. (ed.), Event-Related Potentials: A Methods Handbook. Cambridge, MA: MIT Press, (3-16).

Perkins, W. (1981). *Measurement and maintenance of fluency.* New York, NY: Elsevier.

Perkins, W. H., Kent, R. D., & Curlee, R. F. (1991). A theory of neuropsycholinguistic function in stuttering. *Journal of Speech and Hearing Research, 34,* 734-752. doi: 0022-4685/91/3404-0734$01.00/0.

Peters, H. F. M., Hulstijn, W., Van Lieshout, P. H. H. M. (2000). Recent developments in speech motor research into stuttering. *Folia Phoniatrica et Logopaedica, 52*(1-3), 103-119. doi: 10.1159/000021518.

Picton, T., Bentin, S., Berg, P., Donchin, E., Hillyard, S., Johnson, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(2), 127-152. doi:10.1017/S0048577200000305.

Postma, A. & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research, 36,* 472-487. doi: 0022-4682/933603-0472$01.00/0.

Postma, A., Kolk, H., & Povel, D.J. (1990). Speech planning and execution in stutterers. *Journal of Fluency Disorders, 15*, 49–59.

Preisendorfer, R. W. & Mobley, C. D. (1998). Anatomy of word and sentence meaning. *Proceedings of the National Academy of Sciences, 95*, 899-905.

Prins, D., Main, V., & Wampler, S. (1997). Lexicalization in adults who stutter. *Journal of Speech, language and Hearing Research, 40,* 373-384.

Richman, M. (1986). Rotation of principal components. *Journal of Climatology, 6*, 293-335.

Rosler, F. & Hahne, A. (1992). Hirnelektrische correlate des sprachverstehens zur psycholinguistischen bedeutung der N400-komponente im EEG. *Sprache und Kognition, 11*, 140-161.

Sohlberg, M. M., Mateer, C. A. (2001). Attention Process Training (APT-I): Attention Process Training manual and attention audio CD stimuli manual. Wake Forest, N.C:.Lash & Associates Publishing, 1-9.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C.C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., Tzeng, O., Arévalo, A.L., Vargha, A.,Butler, A.C., Buffington R., & Bates, E. (2004). A new on-line resource for psycholinguistic studies, *Journal of Memory and Language, 51*, 247–250.

Tataryn, D.J., Wood, J. M., & Gorsuch, R. L. (1999). Setting the value of k in promax: A Monte Carlo study. *Educational and Psychological Measurement, 59*, 384-391.

Van der Merwe, A. (1997). *A theoretical framework for the characterization of pathological speech sensorimotor control*. In M.R. McNeil (Ed.), Clinical Management of Sensorimotor Speech Disorders, (1-25). New York, NY:Thieme.

Van Dyke, S. A., Zuverza, V., Hill, L. A., Miller, J. B., Rapport, L. J., Whitman, R. D. (2010). Gender differences in lateralized semantic priming. *Developmental Neuropsychology 34*(4), 381-392. Doi: 10.1080/87565640902964516.

Van Petten, C. & Kutas, M. (1991*). Electrophysiological evidence for the flexibility of lexical processing.* In: Simpson, G. B., (Ed.). Understanding word and sentence. Amsterdam: Elsevier, 129-174.

Viswanath, N., Pointdexter, M., & Rosenfield, D. (1999). Overt phonological errors during part-word repetitions: Some theoretical implications. *Journal of Fluency Disorders*, *24*(2), 107-117. doi:10.1016/S0094-730X(99)00002-9.

Watson, B., Freeman, F., Devous, M., Chapman, S., Finitzo, T., & Pool, K. (1994). Linguistic performance and regional cerebral blood flow in persons who stutter. *Journal of Speech and Hearing Research, 37,* 1221-1228.

Weber-Fox, C. (2001). Neural systems for sentence processing in stuttering. *Journal of Speech, Language, and Hearing Research, 44*, 814–82.

Weber- Fox, C., & Hampton, A. (2008) Stuttering and natural speech processing of semantic and syntactic constraints on verbs. *Journal of Speech, Language, and Hearing Research, 51*, 1058-1071.

Weber-Fox, C., Spencer, R.M.C., Spruill, J.E., & Smith, A. (2004). Phonological processing in adults who stutter: Electrophysiological and behavioral evidence. *Journal of Speech, Language, & Hearing Research, 47,* 1244-1258.

Weiss, A. (2004a). What child language research may contribute to the understanding and treatment of stuttering. *Language, Speech, and Hearing Services in Schools*, *35*(1), 90-92. doi:10.1044/0161-1461(2004/010).

Weiss, A. (2004b). Why we should consider pragmatics when planning treatment for children who stutter. *Language, Speech, and Hearing Services in Schools*, *35*(1), 34-45. doi:10.1044/0161-1461(2004/005).

Wertsch, J.V. (1993). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, MA: Harvard University Press.

World Health Organization (WHO) International Classification of Functioning (2001). *WHO. International Classification of Functioning, Disability, and Health*, p. 18.

Wingate, M. E. (1964). A standard definition of stuttering. *Journal of Speech and Hearing Disorders*, *29*(4), 484-488.

Wingate, M. E. (1988). *The structure of stuttering: A psycholinguistic approach*. New York: Springer-Verlag.

Yaruss, J. S. (2007). Application of the ICF in fluency disorders. *Seminars in Speech and Language, 28*(4), 312-322. doi: 10.1055/s-2007-986528.

Yairi, E. (2007). Subtyping stuttering I: A review. *Journal of Fluency Disorders*, *32*(3), 165-196. doi:10.1016/j.jfludis.2007.04.001.

Appendix

Appendix

picture name	***FR***	***# Syll***	***# Phon***	***best semantic cue (from Nelson & McEvoy)***	***FR***	***# Syll***	***# Phon***	***Check***	***FR***
baby	7	2	4	child	7	1	4	mild	7
basket	7	2	6	weave	7	1	3	leave	7
bed	7	1	3	sleep	7	1	4	weep	7
bell	7	1	3	ring	7	1	3	sing	7
bone	7	1	3	dog	7	1	3		
book	6.91	1	3	read	6.83	1	3	feed	7
bowl		1	3	dish	7	1	3	fish	7
bridge	6.95	1	4	water	7	2	4		
broom	6.92	1	4	sweep	6.93	1	4	keep	7
brush	7	1	4	hair	7	1	2	bear	7
bus	7	1	3	car	7	1	2		
camel	7	2	5	desert	7	2	5		
cat	7	1	3	dog	7	1	3	log	6.73
cheese	7	1	3	mouse	7	1	3	house	7
chest	7	1	4	hair	7	1	2		
clock	7	1	4	time	6.75	1	3	lime	6.92
comb	7	1	3	brush	7	1	4		
desk	6.92	1	4	chair	7	1	2		
dragon	7	2	6	fire	7	1	3	hire	7
fish	7	1	3	water	7	2	3		
flower	7	2	4	rose	6.83	1	3	pose	6.92
glove	7	1	4	hand	7	1	4		
goat	7	1	3	milk	6.92	1	4		
hammock	6.92	2	5	swing	7	1	4	fling	6.83
hand	7	1	4	finger	7	2	4		
hook	6.75	1	3	fish	7	1	3		
hose	6.92	1	3	water	7	2	4	daughter	7
map	7	1	3	road	7	1	3		
match	7	1	3	fire	7	1	3		
monkey	7	2	5	ape	7	1	2	cape	6.92

Appendix (Continued)

picture name	***FR***	***# Syll***	***# Phon***	***best semantic cue (from Nelson & McEvoy)***	***FR***	***# Syll***	***# Phon***	***Check***	***FR***
mountain	7	2	6	hill	7	1	3		
mouse	7	1	3	cat	7	1	3	hat	7
nail	7	1	3	hammer	7	2	4		
net	6.92	1	3	fish	7	1	3		
nose	7	1	3	face	7	1	3	case	6.75
pipe	7	1	3	smoke	7	1	4	joke	7
pizza	7	2	5	food	7	1	3	mood	7
plate		1	4	dish	7	1	3		
popcorn	7	2	7	movie	7	2	4		
priest	7	1	5	church	7	1	3	perch	6.75
rake	7	1	3	leaves	7	1	4		
rock	6.67	1	3	stone	7	1	4	phone	7
sandwich	7	2	7	food	7	1	3		
scissors	7	2	5	cut	7	1	3	hut	7
shoe	6.92	1	2	foot	7	1	3		
shovel	7	2	5	dirt	7	1	3	hurt	7
spider	7	2	5	web	7	1	3		
spoon	7	1	4	fork	7	1	3	cork	7
tent	7	1	4	camp	7	1	4		
thumb	7	1	3	hand	7	1	4		
tie		1	2	neck	7	1	3	deck	7
toilet	7	2	5	bathroom	7	2	6		
whale	7	1	3	fish	7	1	3		
window	7	2	5	glass	7	1	4	bass	7

Table A1. Familiarity ratings (FR), word length in syllables (# syll), and word length in phonemes (# phon) for each target picture label and associated probe word (best semantic cue). Also shown is the lexical item used for the word check on each trial and its familiarity rating.

Appendix (Continued)

picture name	*FR*	*# Syll*	*# Phon*	*phonological cue*	*FR*	*# Syll*	*# Phon*	*Check*	*FR*
baby	7	2	4	bailiff	6.75	2	5		
basket	7	2	6	bashful	7	2	6		
bed	7	1	3	beg	7	1	3	keg	7
bell	7	1	3	bet	7	1	3	met	6.83
bone	7	1	3	boulder	6.92	2	5	older	7
book	6.91	1	3	bushel	7	2	5		
bowl		1	3	boast	6.92	1	4	roast	6.5
bridge	6.95	1	4	brim		1	4	trim	6.83
broom	6.92	1	4	bruise	7	1	4		
brush	7	1	4	brunt	5	1	5	grunt	7
bus	7	1	3	bunch	7	1	4	hunch	6.67
camel	7	2	5	candid	6.83	2	6		
cat	7	1	3	cab	7	1	3	tab	7
cheese	7	1	3	cheer	7	1	3	hear	7
chest	7	1	4	checker	6.92	2	4		
clock	7	1	4	clot	7	1	4	pot	7
comb	7	1	3	cold	7	1	4	mold	7
desk	6.92	1	4	debit	5.75	2	5		
dragon	7	2	6	draft	6.83	1	5	craft	7
fish	7	1	3	fiddle	6.92	2	5	middle	6.58
flower	7	2	4	flounder	6.83	2	6		
glove	7	1	4	glutton	6.67	2	6		
goat	7	1	3	gold	7	1	4	cold	7
hammock	6.92	2	5	handle	7	2	6		
hand	7	1	4	hamper	7	2	5	damper	7
hook	6.75	1	3	hoof	6.5	1	3		
hose	6.92	1	3	hopeful	6.92	2	6		
map	7	1	3	mask	7	1	4	flask	6.92
match	7	1	3	magnet	6.75	2	6		
monkey	7	2	5	mustard	7	2	6		
mountain	7	2	6	mouth	7	1	3	south	7
mouse	7	1	3	mound	6.58	1	4		
nail	7	1	3	neighbor	7	2	4		
net	6.92	1	3	necklace	7	2	6		
nose	7	1	3	note	7	1	3	float	7
pipe	7	1	3	pine	6.92	1	3	spine	7
pizza	7	2	5	peacock	6.83	2	5		
plate		1	4	plague	6.75	1	4		
popcorn	7	2	7	posture	7	2	6		
priest	7	1	5	preview	7	2	6		
rake	7	1	3	racial	7	2	5	facial	6.83
rock	6.67	1	3	romp	7	1	4		

Appendix (Continued)

picture name	*FR*	*# Syll*	*# Phon*	*phonological cue*	*FR*	*# Syll*	*# Phon*	*Check*	*FR*
scissors	7	2	5	simple	7	2	6	dimple	6.92
shoe	6.92	1	2	shoots		1	4	boots	7
shovel	7	2	5	shutter	7	2	4	flutter	6.67
spider	7	2	5	spike	7	1	4	like	7
spoon	7	1	4	spooky	7	2	5		
tent	7	1	4	temper	7	2	5		
thumb	7	1	3	thunder	7	2	5		
tie		1	2	tile	6.92	1	3		
toilet	7	2	5	toys	7	1	4	boys	7
whale	7	1	3	wait	7	1	3	late	6.75
window	7	2	5	whimper	6.75	2	5		

Table A2. Familiarity ratings (FR), word length in syllables (# syll), and word length in phonemes (# phon) for each target picture label and associated probe word (phonological cue). Also shown is the lexical item used for the word check on each trial and its familiarity rating.

picture name	*FR*	*# Syll*	*# Phon*	*Control set for Semantic Cue*	*Check*	*FR*	*Control Set for Phonolog-ical Cue*	*Check*	*FR*
baby	7	2	4	sweep			note		
basket	7	2	6	ring			mound	*sound*	7
bed	7	1	3	cut			draft		
bell	7	1	3	hair	fair	7	spike		
bone	7	1	3	time			checker	*wrecker*	7
book	6.91	1	3	water			thunder	*plunder*	5.33
bowl		1	3	fire	liar	6.92	peacock		
bridge	6.95	1	4	dog	frog	7	pine		
broom	6.92	1	4	child			hopeful		
brush	7	1	4	mouse			temper		
bus	7	1	3	hair			simple		
camel	7	2	5	dish			mask		
cat	7	1	3	rose			racial		
cheese	7	1	3	hand	land	6.92	romp	*chomp*	
chest	7	1	4	car	far	6.58	whimper		
clock	7	1	4	water			brim		
comb	7	1	3	fish	wish	6.92	shoots		
desk	6.92	1	4	water			bunch		
dragon	7	2	6	chair	stair	7	bet		
fish	7	1	3	road	load	6.7	preview		

Appendix (Continued)

picture name	***FR***	***# Syll***	***# Phon***	***Control set for Semantic Cue***	***Check***	***FR***	***Control Set for Phonolog-ical Cue***	***Check***	***FR***
flower	7	2	4	weave			spooky	*kooky*	
glove	7	1	4	ape			neighbor	*labor*	7
goat	7	1	3	leaves	sleeves	7	shutter		
hammock	6.92	2	5	desert			plague	vague	7
hand	7	1	4	sleep			boast		
hook	6.75	1	3	swing			sandal		
hose	6.92	1	3	food			glutton	mutton	
map	7	1	3	fish			flounder	pounder	7
match	7	1	3	fish			hoof		
monkey	7	2	5	hill	bill	7	debit		
mountain	7	2	6	finger	linger	6.67	toys		
mouse	7	1	3	fire			handle	sandal	6.58
nail	7	1	3	smoke			brunt		
net	6.92	1	3	face			mustard	custard	6.75
nose	7	1	3	read			wait		
pipe	7	1	3	cat			bashful		
pizza	7	2	5	dog			necklace		
plate		1	4	movie	groovy	7	bushel		
popcorn	7	2	7	church			bruise	cruise	7
priest	7	1	5	food			tile	pile	7
rake	7	1	3	bathroom			cold		
rock	6.67	1	3	foot			fiddle		
sandwich	7	2	7	hammer	slammer	6.92	posture		
scissors	7	2	5	milk			candid	landed	6.58
shoe	6.92	1	2	stone			bailiff		
shovel	7	2	5	neck			cab		
spider	7	2	5	hand			magnet		
spoon	7	1	4	web			clot		
tent	7	1	4	dish			gold		
thumb	7	1	3	glass			boulder		
tie		1	2	dirt			mouth		
toilet	7	2	5	fork			beg		
whale	7	1	3	camp			hamper		
window	7	2	5	brush	crush	7	cheer		

Table A3. Familiarity ratings (FR), word length in syllables (# syll), and word length in phonemes (# phon) for each target picture label and unrelated probe word. Also shown is the lexical item used for the word check on each trial and its familiarity rating.